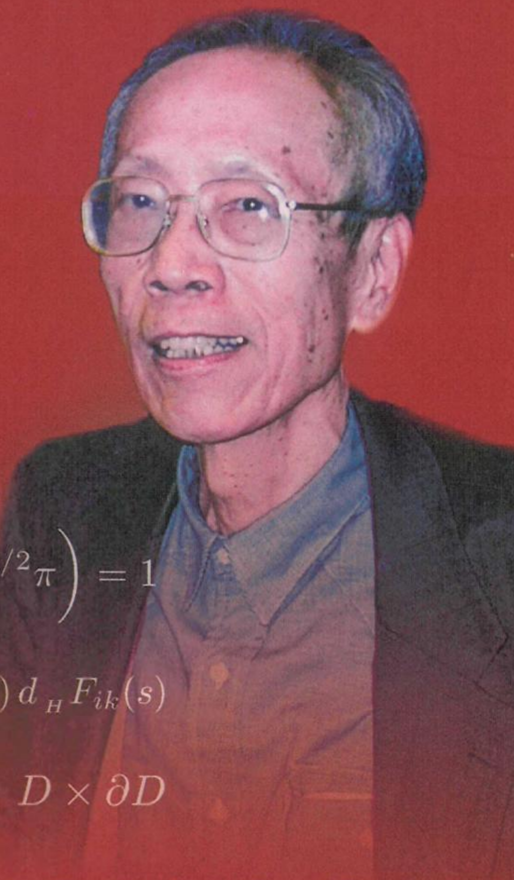


$$\Pr\left(\liminf \frac{S_n^*}{s_n(\lg_2 s_n)^{-1/2}} = 8^{-1/2}\pi\right) = 1$$

$${}_H p_{ij}(t) = {}_{k,H}p_{ij}(t) + \int_0^t {}_H p_{kj}(t-s)\, d\, {}_H F_{ik}(s)$$

$$u(x,z) = E_z^x\left\{e_q(\tau_D)\right\} \neq \infty \text{ in } D \times \partial D$$

SELECTED WORKS OF KAI LAI CHUNG

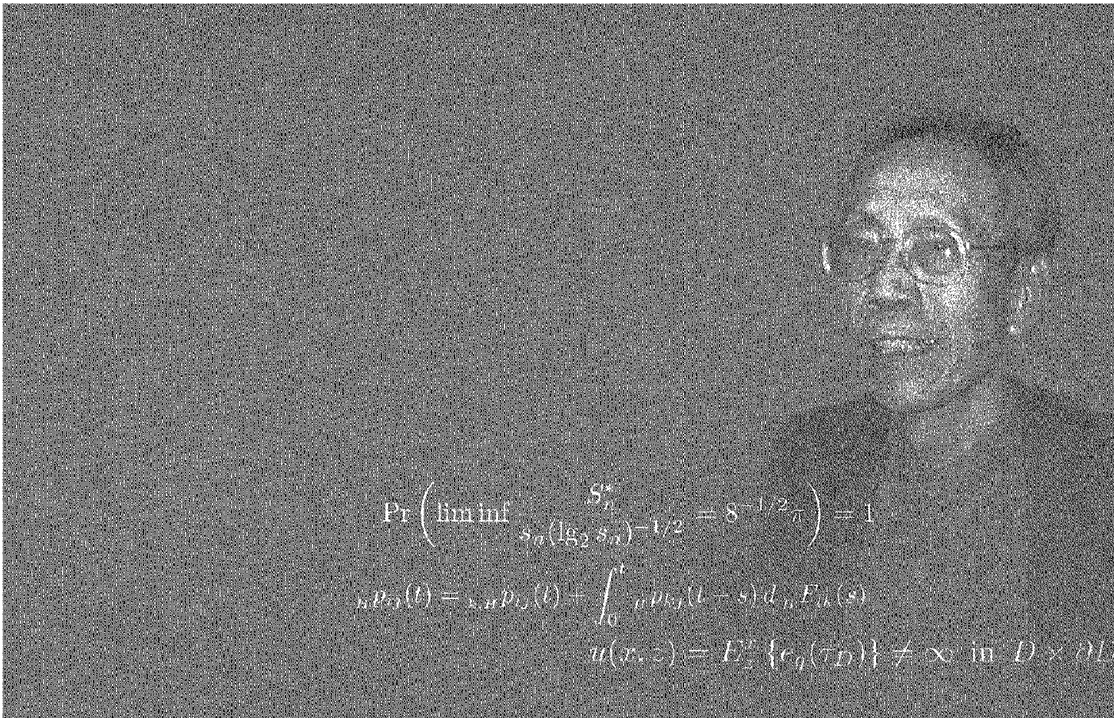
Editors

Farid AitSahlia • Elton Hsu • Ruth Williams

World Scientific

SELECTED WORKS OF
KAI LAI CHUNG

This page intentionally left blank



SELECTED WORKS OF KAI LAI CHUNG

Editors

Farid AitSahlia
University of Florida, USA

Elton Hsu
Northwestern University, USA

Ruth Williams
University of California-San Diego, USA

World Scientific

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI

Published by

World Scientific Publishing Co. Pte. Ltd.
5 Toh Tuck Link, Singapore 596224
USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601
UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

SELECTED WORKS OF KAI LAI CHUNG

ISBN-13 978-981-283-385-3
ISBN-10 981-283-385-4

Printed in Singapore by B & JO Enterprise

Preface

We are pleased to present this volume of selected works in celebration of the 90th birthday of Professor Kai Lai Chung.

Kai Lai Chung was born in 1917 in Shanghai, China. His family home though was in Hangzhou in Zhejiang Province. He entered Tsinghua University in 1936 and first studied physics but graduated in mathematics in 1940. During the war with Japan, major universities in the Beijing-Tianjin region moved to the southwest city of Kunming and regrouped as the National Southwestern Associated University and Chung worked there in a position analogous to that of assistant professor. During this period, he first studied number theory with Lo-Keng Hua and then probability theory with Pao-Lu Hsu. In 1944, Kai Lai Chung won a highly competitive Boxer Rebellion Indemnity scholarship for study in the United States and he arrived at Princeton University in December, 1945. He completed his Ph.D. at Princeton in 1947 with Harold Cramér as advisor (Cramér was visiting Princeton at the time — S. Wilks and J. Tukey were the other members of the dissertation committee). Chung's thesis was entitled "On the maximum partial sum of sequences of independent random variables". Kai Lai Chung subsequently held academic appointments at the University of Chicago, Columbia University, University of California at Berkeley, Cornell University and Syracuse University. He moved to Stanford University in 1961 and is currently emeritus Professor of Mathematics at Stanford. Over the years, he held extended visiting appointments at several institutions: University of Strasbourg (France), University of Pisa (Italy), and the ETH (Eidgenössische Technische Hochschule) of Zurich (Switzerland). He held the George A. Miller Visiting Professorship at the University of Illinois at Urbana-Champaign in 1970-71.

Kai Lai Chung taught probability for over 30 years and supervised 14 Ph.D. students. The Mathematics Genealogy project currently lists a total of 112 academic descendants for him. The Ph.D. students (in chronological order) were as follows: Warren Hirsch, Rafael Chacon, William Pruitt, Norman Pullman, Naresh Jain, Arthur Pittenger, Robert Smythe, Michael Chamberlain, Christopher Nevison, Michael Steele, Ruth Williams, Elton Hsu, Ming Liao, Vassilis Papanicolaou.

This volume contains a selection of Kai Lai Chung's extensive journal publications, which span a period of 70 years. The selection was made in consultation with him and is only a subset of the many contributions that he made throughout his prolific career. Another volume, *Chance and Choice*, published by World Scientific in 2004, contains yet another subset, with four articles in common with this volume. Kai Lai Chung's research contributions had a major influence on several areas in probability. Among his most significant works are those related to sums of independent random variables, Markov chains, time reversal of Markov processes, probabilistic potential theory, Brownian excursions, and gauge theorems for the Schrödinger equation. We have included commentary articles by Naresh Jain, Ronald Getoor, Ruth Williams and Michael Cranston elaborating on Kai Lai Chung's contributions in these areas and on the further developments that they spawned.

In 1981, Kai Lai Chung, along with Ronald Getoor and Erhan Çinlar, initiated the "Seminars on Stochastic Processes". These conferences, with their innovative structure of just a few formal talks, allowing plenty of time for informal discussions and research problem sessions, continue as highly successful annual meetings to this day. We are pleased to include here several of Kai Lai Chung's articles from proceedings volumes for the "Seminar", published by Birkhäuser. In addition to his research articles, Kai Lai Chung's eleven books have influenced generations of students of probability, both graduate and undergraduate. He is well known for his clear, precise and entertaining style.

Kai Lai Chung played an influential role in the development of probability theory in his native China immediately after the chaotic years of the Cultural Revolution (1966-1976). His visit to China in 1978 (together with J. L. Doob and J. Neveu) was the starting point for renewed contact of Chinese probabilists with the West. He has visited China many times since then, given numerous lectures and short courses, and helped young Chinese students gain opportunities to study in the United States.

We thank Naresh Jain, Ronald Getoor and Michael Cranston for contributing commentary articles for this volume. We are indebted to Erhan Çinlar, Rafe Mazzeo and Renming Song for their help in providing biographical information and to Louis Chen and Tze Leung Lai for their assistance in making connection with World Scientific. We are grateful to the various publishers for permission to reprint the articles that are included here, and also to Yubing Zhai and Ji Zhang of World Scientific for assistance in preparing this volume. We also wish to thank Lilia and Marilda Chung for providing a photograph of Kai Lai. Finally, we express our thanks to Kai Lai for his inspiration and guidance, and his many engaging conversations and stimulating questions over the years.

Farid AitSahlia, Gainesville, FL
Elton P. Hsu, Evanston, IL
Ruth J. Williams, La Jolla, CA
April 30, 2008

Contents

Chung Kai Lai

PART 1

Bibliography

This page intentionally left blank

Bibliography of Kai Lai Chung

Articles

* = included in the selection

[1936a] Huoranlü dingli zhi chudeng zhengming (Elementary proof of a theorem in probability theory) (in Chinese, with Bingzhen Liu), *Mathematical Magazine* **1**, no. 4 (1936).

[1940a] Two remarks on Viggo Brun's method, *Sci. Rep. Nat. Tsing Hua Univ., Kuming, Ser. A* **4** (1940), 249–255.

[1940b] Sur un théorème de M. Gumbel (in French), *C. R. Acad. Sci., Paris* **210** (1940), 620–621.

[1941a] Note on a theorem on quadratic residues, *Bull. Amer. Math. Soc.* **47** (1941), 514–516.

[1941b] A generalization of an inequality in the elementary theory of numbers, *J. Reine Angew. Math.* **183** (1941), 193–196.

*[1941c] On the probability of the occurence of at least m events among n arbitrary events, *Ann. Math. Stat.* **12** (1941), 328–338.

*[1942a] On mutually favorable events, *Ann. Math. Stat.* **13** (1942), 338–349.

[1943a] Generalization of Poincaré's formula in the theory of probability, *Ann. Math. Stat.* **14** (1943), 63–65.

[1943b] On fundamental systems of probabilities of a finite number of events, *Ann. Math. Stat.* **14** (1943), 123–133.

[1943c] Further results on probabilities of a finite number of events, *Ann. Math. Stat.* **14** (1943), 234–237.

[1945a] A combinatorial formula and its application to the theory of probability of arbitrary events (with Lietz C. Hsu), *Ann. Math. Stat.* **16** (1945), 91–95.

[1946a] The approximate distribution of Student's statistic, *Ann. Math. Stat.* **17** (1946), 447–465.

[1946b] Sur un théorème de probabilités dénombrables (in French, with P. L. Hsu), *C. R. Acad. Sci. Paris* **223** (1946), 467–469.

[1947a] On the maximum partial sum of independent random variables, *Proc. Nat. Acad. Sci. U. S. A.* **33** (1947), 132–136.

[1947b] Note on some strong laws of large numbers, *Amer. J. Math.* **69** (1947), 189–192.

*[1947c] On the lower limit of sums of independent random variables (with Paul Erdös), *Ann. Math. (2)* **48** (1947), 1003–1013.

[1948a] Asymptotic distribution of the maximum cumulative sum of independent random variables, *Bull. Amer. Math. Soc.* **54** (1948), 1162–1170.

*[1948b] On the maximum partial sums of sequences of independent random variables, *Trans. Amer. Math. Soc.* **64** (1948), 205–233.

[1948c] On a lemma by Kolmogoroff, *Ann. Math. Stat.* **19** (1948), 88–91.

*[1949a] On the zeros of $\sum_1^n \pm 1$ (with G. A. Hunt), *Ann. Math. (2)* **50** (1949), 385–400.

[1949b] An estimate concerning the Kolmogoroff limit distribution, *Trans. Amer. Math. Soc.* **67** (1949), 36–50.

[1949c] On fluctuations in coin-tossing (with W. Feller), *Proc. Nat. Acad. Sci. U.S.A.* **35** (1949), 605–608.

*[1950a] Fluctuations of sums of independent random variables, *Ann. Math. (2)* **51** (1950), 697–706. Correction: **57** (1953), 604–605.

[1950b] An ergodic theorem for stationary Markov chains with a countable number of states, *Proceedings of Intern. Congr. Math., Cambridge, MA (1950)*, AMS (1952), p.568.

*[1951a] The strong law of large numbers, *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950*, University of California Press, Berkeley and Los Angeles (1951), 341–352.

*[1951b] On the distribution of values of sums of random variables (with W. H. J. Fuchs), in *Four Papers on Probability Theory, Mem. Amer. Math. Soc.* **6** (1951), 12 pp.

*[1951c] Probability limit theorems assuming only the first moment, **I** (with P. Erdös), in *Four Papers on Probability Theory, Mem. Amer. Math. Soc.* **6** (1951), 19 pp.

[1951d] Remarks on fluctuations of sums of independent random variables (with M. Kac), *Mem. Amer. Math. Soc.* **6** (1951), 11 pp. Correction: *Proc. Amer. Math. Soc.* **4** (1953), 560–563.

[1952a] An extension of renewal theory (with Harry Pollard), *Proc. Amer. Math. Soc.* **3** (1952), 303–309.

*[1952b] On the application of the Borel-Cantelli lemma (with P. Erdös), *Trans. Amer. Math. Soc.* **72** (1952), 179–186.

[1952c] On a limit theorem in renewal theory (with J. Wolfowitz), *Ann. Math. (2)* **55** (1952), 1–6.

[1952d] On the renewal theorem in higher dimensions, *Skand. Aktuarietidskr* **35** (1952), 188–194.

*[1953a] Contributions to the theory of Markov chains, *J. Research Nat. Bur. Standards* **50** (1953), 203-208.

[1953b] Sur les lois de probabilité unimodales (in French), *C. R. Acad. Sci. Paris* **236** (1953), 583–584.

[1954a] On a stochastic approximation method, *Ann. Math. Stat.* **25** (1954), 463–483.

*[1954b] Contributions to the theory of Markov chains, II, *Trans. Amer. Math. Soc.* **76** (1954), 397–419.

[1956a] Foundations of the theory of continuous parameter Markov chains, in *Proc. Third Berkeley Sympos. Math. Stat. and Prob., II*, Univ. of California Press, Berkeley, CA (1956), 29–40.

[1956b] Non-recurrent random walks (with C. Derman), *Pacific J. Math* **6** (1956), 441–447.

*[1956c] Some new developments in Markov chains, *Trans. Amer. Math. Soc.* **81** (1956), 195–210.

[1957a] Some aspects of continuous parameter Markov chains, *Publ. Inst. Stat. Univ. Paris* **6** (1957), 271–287.

*[1958a] On a basic property of Markov chains, *Ann. Math. (2)* **68** (1958), 126–149.

*[1958b] Continuous parameter Markov chains, invited address, *Proceedings of Intern. Congr. Math.*, Edinburgh (1958), 510–517.

*[1959a] On the Lipschitz's condition for Brownian motion (with P. Erdös and T. Sirao), *J. Math. Soc. Japan* **11** (1959), 263–274.

*[1960a] On last exit times, *Illinois J. Math.* **4** (1960), 629–639.

*[1961a] Some remarks on taboo probabilities, *Illinois J. Math.* **5** (1961), 431–435.

*[1961b] Probabilistic methods in Markov chains, in *Proc. 4th Berkeley Sympos. Math. Stat. and Prob., II*, Univ. of California Press, Berkeley, CA (1961), 35–56.

[1961c] A note on the ergodic theorem of information theory, *Ann. Math. Stat.* **32** (1961), 612–614.

[1962a] On the exponential formulas of semi-group theory, *Math. Scand.* **10** (1962), 153–162.

[1962b] On the Martin boundary for Markov chains, *Proc. Nat. Acad. Sci. U.S.A.*

48 (1962), 963–968.

*[1962c] On the recurrence of sums of random variables (with Donald Ornstein), *Bull. Amer. Math. Soc.* **68** (1962), 30–32.

[1962d] The ergodic theorem of information theory, in *Recent Developments in Information and Decision Processes*, MacMillan (1962), 141–148.

*[1963a] On the boundary theory for Markov chains, *Acta Math.* **110** (1963), 19–77.

*[1964a] The general theory of Markov processes according to Doeblin, *Z. Wahrsch. Verw. Gebiete* **2** (1964), 230–254.

[1965a] Sur une équation de convolution (in French), *C. R. Acad. Sci. Paris* **260** (1965), 4665–4667.

[1965b] Sur une équation de convolution (in French), *C. R. Acad. Sci. Paris* **260** (1965), 6794–6796.

*[1965c] Fields, optionality and measurability (with J. L. Doob), *Amer. J. Math.* **87** (1965), 397–424.

*[1966a] On the boundary theory for Markov chains, II, *Acta Math.* **115** (1966), 111–163.

[1967a] Markov processes with infinities, *Markov Processes and Potential Theory (Proc. Sympos. Math. Res. Center, Madison, Wis., 1967)*, Wiley, New York (1967), 23–32.

*[1969a] To reverse a Markov process (with John B. Walsh), *Acta Math.* **123** (1969), 225–251.

[1970a] On diverse questions of time reversing in Markov chains (and processes), *Proceedings of the Twelfth Biennial Seminar of the Canadian Mathematical Congress on Time Series and Stochastic Processes; Convexity and Combinatorics (Vancouver, B.C., 1969)*, Canad. Math. Congr., Montreal, Que. (1970), 165–175.

*[1971a] Boundary behavior of Markov chains and its contributions to general processes, invited address, *Actes du Congrès International des Mathématiciens (Nice, 1970), Tome 2* (1971), 499–505.

[1971b] A simple proof of Doob's convergence theorem, *Séminaire de Probabilités, V, 1969/1970*, Lecture Notes in Math. **191**, Springer-Verlag, Berlin (1971), p. 76.

[1972a] Some universal field equations, in *Séminaire de Probabilités, VI, 1970/1971*, Lecture Notes in Math **258**, Springer-Verlag, Berlin (1972), 90–97. Russian translation: *Matematika (Moskva)* **17**, no. 6 (1973), 49-52.

*[1972b] Crudely stationary counting processes, *Amer. Math. Monthly* **79** (1972), 867–877.

*[1972c] On the fundamental hypotheses of Hunt processes, *Symposia Mathematica*, Vol. IX (Convegno di Calcolo delle probabilità, INDAM, Rome, 1971),

Academic Press, London (1972), 43–52.

[1972d] The Poisson process as renewal process, *Period. Math. Hungar.* **2** (1972), 41–48.

*[1973a] Probabilistic approach in potential theory to the equilibrium problem, *Ann. Inst. Fourier (Grenoble)* **23**, no. 3 (1973), 313–322.

[1973b] An expression for canonical entrance laws, in *Stochastic Analysis (A Tribute to the Memory of Rollo Davidson)*, Wiley, London (1973), 311–314.

*[1974a] Meyer's theorem on predictability (with John B. Walsh), *Z. Wahrsch. Verw. Gebiete* **29** (1974), 253–256.

[1974b] Note on last exit decomposition, in *Séminaire de Probabilités, VIII, 1972/1973*, Lect. Notes in Math **381**, Springer-Verlag, Berlin (1974), 20–21.

[1975a] Maxima in Brownian excursions, *Bull. Amer. Math. Soc.* **81** (1975), 742–745.

*[1975b] Approximation of a continuous parameter Markov chain by discrete skeletons, *Bull. London Math. Soc.* **7** (1975), 71–76.

*[1975c] Remarks on equilibrium potential and energy. Collection of articles dedicated to Marcel Brelot on the occasion of his 70th birthday, *Ann. Inst. Fourier (Grenoble)* **25**, no. 3-4 (1975), 131–138.

[1975d] A bivariate distribution in regeneration, *J. Appl. Prob.* **12**, no. 4 (1975), 837–839.

[1975e] Emptiness times of a dam with stable input and general release function (with P. J. Brockwell), *J. Appl. Prob.* **12** (1975), 212–217.

*[1976a] Downcrossings and local time (with Richard Durrett), *Z. Wahrsch. Verw. Gebiete* **35**, no. 2 (1976), 147–149.

[1976b] A proof of a lemma of A. V. Skorohod (in Russian), *Teor. Verojatnost. i Mat. Stat.* **15** (1976), 151–152, 157. English translation: *Theory Probab. Math. Stat.* **15** (1978), p. 156.

*[1976c] Excursions in Brownian motion, *Ark. Mat.* **14**, no. 2 (1976), 155–177.

[1977a] Pedagogic notes on the barrier theorem, in *Séminaire de Probabilités, XI, 1975/1976*, Lecture Notes in Math. **581**, Springer-Verlag, Berlin (1977), 27–33. Correction: in *Séminaire de Probabilités, XII, 1976/1977*, Lecture Notes in Math. **649**, Springer-Verlag, Berlin (1978), p. 739.

*[1977b] The condenser problem (with R. K. Getoor), *Ann. Prob.* **5**, no. 1 (1977), 82–86.

*[1979a] Left continuous moderate Markov processes (with Joseph Glover), *Z. Wahrsch. Verw. Gebiete* **49**, no. 3 (1979), 237–248.

[1979b] Pao Lu Hsü: 1909–1970 (with T. W. Anderson and E. L. Lehmann), *Ann.*

Stat. **7**, no. 3 (1979), 467-470. Chinese translation: *Knowledge Practice in Math.* **3** (1980), 3–5.

[1979c] Hsü's work in probability, *Ann. Stat.* **7**, no. 3 (1979), 479–483. Chinese translation: *Knowledge and Practice in Math.* **3** (1980), 12–15.

*[1980a] Equilibrium and energy (with K. Murali Rao), *Probab. Math. Stat.* **1**, no. 2 (1980), 99–108.

*[1980b] Kac functional and Schrödinger equation (with S. R. S. Varadhan), *Studia Math.* **68**, no. 3 (1980), 249–260.

[1980c] On recurrence of a random walk in the plane (with Torgny Lindvall), *Proc. Amer. Math. Soc.* **78**, no. 2 (1980), 285–287.

*[1980d] On stopped Feynman-Kac functionals, in *Séminaire de Probabilités, XIV, 1978/1979*, Lecture Notes in Math **784**, Springer-Verlag, Berlin (1980), 347–356.

[1980e] Sur la théorie du potentiel avec la fonctionelle de Feynman-Kac (in French) (with K. Murali Rao), *C. R. Acad. Sci. Paris Sér. A-B* **290**, no. 13 (1980), A629–A631.

*[1980f] A new setting for potential theory. I (with K. Murali Rao), *Ann. Inst. Fourier (Grenoble)* **30**, no. 3 (1980), 167–198. Erratum: **31**, no. 2 (1981), loose sheet.

*[1981a] Feynman-Kac functional and the Schrödinger equation (with K. Murali Rao), in *Seminar on Stochastic Processes (Evanston, IL, 1981)*, Progress in Probab. Stat. **1**, Birkhäuser, Boston (1981), 1–29.

[1981b] Brownian motion on the line, *J. Math. Res. Exposition* **1** (1981), 73–82.

[1982a] Brownian motion on the line. II, *J. Math. Res. Exposition* **2**, no. 3 (1982), 87–98.

[1982b] A cluster of great formulas, *Acta Math. Acad. Sci. Hungar.* **39**, no. 1-3 (1982), 65–67.

[1982c] Ping erze kexue tongxun (Review of two scientific communications) (in Chinese), *J. Math. Res. Expos.*, **2**, no. 3 (1982), 141–142.

[1983a] An inequality for boundary value problems, in *Seminar on Stochastic Processes (Evanston, IL, 1982)*, 111–122, Progress in Probab. Stat. **5**, Birkhäuser, Boston (1983).

*[1983b] Extension of domains with finite gauge (with R. Durrett and Zhongxin Zhao), *Math. Ann.* **264**, no. 1 (1983), 73–79.

[1983c] Properties of finite gauge with an application to local time, in *Probability and Mathematical Statistics, Essays in honour of Carl-Gustav Esseen*, edited by Allan Gut and Lars Holst, Uppsala Univ., Uppsala (1983), 16–24.

[1984a] Conditional gauges, in *Seminar on Stochastic Processes (Gainesville, FL, 1983)*, Progress in Probab. Stat. **7**, Birkhäuser, Boston (1984), 17–22.

*[1984b] The lifetime of conditional Brownian motion in the plane, *Ann. Inst. H. Poincaré Probab. Stat.* **20**, no. 4 (1984), 349–351.

*[1984c] Duality under a new setting (with Ming Liao and K. M. Rao), in *Seminar on Stochastic Processes (Gainesville, FL, 1983)*, Progress in Probab. Stat. **7**, Birkhäuser, Boston (1984), 23–38.

[1984d] Some probabilistic results on the Sturm-Liouville equations (with Zhongxin Zhao), *J. Systems Sci. Math. Sci.* **4**, no. 4 (1984), 320–324.

[1984e] Mathematics and applications, *Math. Medley* **12**, no. 1 (1984), 1–8. Chinese translation: *Ziran Zazhi (Nature Magazine)* **11**, no. 9 (1988).

*[1985a] The gauge and conditional gauge theorem, in *Séminaire de Probabilités, XIX, 1983/1984*, Lecture Notes in Math **1123**, Springer-Verlag, Berlin (1985), 496–503.

[1985b] A probability anecdote, *J. Math. Res. Exposition* **5**, no. 2 (1985), 141–143.

[1985c] A simple proof of a new theorem on conditional Brownian motion (in Chinese), *Kexue Tongba (Science Reports)* **30**, no. 5 (1985), 329–332.

[1985d] Probabilistic approach to boundary value problems for Schrödinger's equation, *Expositiones Math.* **3**, no. 2 (1985), 175–178.

[1985e] Remark on the conditional gauge theorem, in *Séminaire de Probabilités, XIX, 1983/1984*, Lecture Notes in Math **1123**, Springer-Verlag, Berlin (1985), 423–425.

[1986a] Notes on the inhomogeneous Schrödinger equation, in *Seminar on Stochastic Processes (Evanston, IL, 1984)*, Progress in Probab. Stat. **9**, Birkhäuser, Boston (1986), 55–62.

*[1986b] Gauge theorem for the Neumann problem (with Pei Hsu), in *Seminar on Stochastic Processes (Evanston, IL, 1984)*, Progress in Probab. Stat. **9**, Birkhäuser, Boston (1986), 63–70.

[1986c] Comparison of probability and eigenvalue methods for the Schrödinger equation (with Peter Li), in *Probability, Statistical Mechanics, and Number Theory, Adv. Math. Suppl. Stud.*, **9**, Academic Press, Orlando, FL (1986), 25–34.

[1986d] Comparison of probability and classical methods for the Schrödinger equation (with P. Li and R. J. Williams), *Exposition. Math.* **4**, no. 3 (1986), 271–278.

*[1986e] Doubly-Feller process with multiplicative functional, in *Seminar on Stochastic Processes (Gainesville, FL, 1985)*, Progress in Probab. Stat. **12**, Birkhäuser, Boston (1986), 63–78.

[1986f] "Markov chain must have a beginning", in memory of Prof. Loo Keng Hua, *J. Math. Res. Exposition (English ed.)* (1986), 1–3.

[1987a] Pólya's work in probability, *Bull. London Math. Soc.* **19** (1987), 559–608.

[1987b] Carleson's proof of a remarkable equation, *J. Math. Res. Exposition* **7**, no. 1 (1987), 117–123.

[1987c] Green's function for a ball, in *Seminar on Stochastic Processes (Charlottesville, VA, 1986)*, Progress in Probab. Stat. **13**, Birkhäuser, Boston (1987), 1–13.

[1988a] Reminiscences of some of Paul Lévy's ideas in Brownian motion and in Markov Chains, in *Colloque Paul Lévy sur les Processus Stochastiques (Palaiseau, 1987), Astérisque*, **157–158** (1988), 29–36.

[1988b] Probability methods in potential theory, in *Potential Theory–Surveys and Problems (Prague, 1987)*, Lect. Notes in Math **1344**, Springer-Verlag, Berlin (1988), 42–54.

*[1988c] General guage theorem for multiplicative functionals (with K. M. Rao), *Trans. Amer. Math. Soc.* **306**, no. 2 (1988), 819–836.

[1989a] Gauge theorem for unbounded domains, in *Seminar on Stochastic Processes (Gainesville, FL, 1988)*, Progress in Probab. Stat. **17**, Birkhäuser, Boston (1989), 87–98.

[1989b] Reminiscences of some of Paul Lévy's ideas in Brownian motion and in Markov chains (with a postscript), in *Seminar on Stochastic Processes (Gainesville, FL, 1988)*, Progress in Probab. Stat. **17**, Birkhäuser, Boston (1989), 99–107.

*[1991a] Paul Lévy's way to his local time (with A. A. Balkema), in *Seminar on Stochastic Processes, Vancouver (1990)*, Birkhäuser, Boston (1991), 5–14.

*[1992a] Greenian bounds for Markov processes, *Potential Anal.* **1**, no. 1 (1992), 83–92.

*[1993a] Doeblin's big limit theorem, *J. Theoret. Probab.* **6**, no. 3 (1993), 417–426.

[1993b] Reminiscences of one of Doeblin's papers, in *Doeblin and Modern Probability (Blaubeuren, 1991), Contemp. Math.* **149**, Amer. Math. Soc., Providence, RI (1993), 65–71.

[1994a] Stochastic analysis of Q-matrix, in *Selected Topics on Stochastic Modelling*, World Scientific Publishing Co. River Edge, NJ (1994), 3–25.

[1998a] Multinomial ratio [Paul Erdös solves a problem], *Methods Appl. Anal.* **5**, no. 2 (1998), 143–156.

[1998b] Probability and Doob, *Amer. Math. Monthly* **105**, no. 1 (1998), 28–35.

[1999a] On the return to equilibrium (in Italian), *Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl.* **10** (1999), 213–218.

[1999b] Borel et la martingale de Saint-Pétersbourg (in French) (with Bernard Bru and Marie-France Bru), *Rev. Histoire Math.* **5**, no. 2 (1999), 181–247.

[2003a] New solution of famous theorem, *J. Math. Res. Exposition* **23**, no. 4 (2003), Inside front cover.

[2004a] Note on open problem, *J. Math. Res. Exposition* **24**, no. 3 (2004), Inside front cover.

[2005a] An unsolved problem by Feller, *J. Math. Res. Exposition* **25**, no. 3 (2005), 463–464.

Books

[1960] *Markov chains with stationary transition probabilities*, Grund. Math. Wiss. **104**, Springer-Verlag, New York/Berlin (1960), x+278 pp.
- second edition (1967), xi+301 pp.
- Russian translation: Однородные Цепи Маркова, translated by V. F. Kolchin, Mir, Moscow (1964)

[1968] *A Course in Probability Theory*, Harcourt, Brace & World, Inc., New York (1968), xii+331 pp.
- second edition, Probability and Mathematical Statistics, Vol. 21, Academic Press (1974), xii+365 pp.
- second revised edition (2000).
- third edition with Supplement (2001), xviii+419 pp.

[1970] *Lectures on boundary theory for Markov chains* (with the cooperation of Paul-André Meyer), Annals of Mathematical Studies, no. 65, Princeton University Press, Princeton, NJ, and University of Tokyo Press, Tokyo (1970), xvii+94 pp.

[1974] *Elementary Probability with Stochastic Processes*, Springer-Verlag, New York/Heidelberg (1974), x+325 pp.
- second edition (1975), x+325 pp.
- third edition (1979), xvi+325 pp.
- fourth edition (2004) (with Farid AitSahlia), includes an Introduction to Mathematical Finance, xvi+402 pp.
- German translation: *Elementare Wahscheinlichkeitstheorie und Stochastische Prozesse*, translated by Herbert Vogt, second edition, Hochschultext, Springer-Verlag, Berlin/New York (1978), ix+343 pp.
- Spanish translation.
- Persian translation.
- Chinese translation.
- Russian translation, 4th edition, 2007.

[1982] *Lectures From Markov Processes to Brownian Motion*, Grund. Math. Wiss. **249**, Springer-Verlag, New York/Berlin (1982), viii+239 pp.
- second edition (2005) (with John B. Walsh), includes Time Symmetry, xii+431 pp.

[1983] *Introduction to Stochastic Integration* (with R. J. Williams), Progress in Probab. Stat. **4**, Birkhäuser, Boston (1983), xiii+191 pp.

– second edition (1990), xvi+276 pp.
– Russian translation: Введение в Стохастическое Интегрирование, translated with a preface by V. M. Kruglov, Mir, Moscow (1987), 152 pp.

[1995a] *Green, Brown, and Probability*, World Scientific, River Edge, NJ (1995), xiv+106 pp.
– second edition (2001), includes Brownian motion on the line, x+170 pp.

[1995b] *From Brownian Motion to Schrödinger's Equation* (with Zhongxin Zhao), Grund. Math. Wiss. **312**, Springer-Verlag, Berlin (1995), xii+287 pp.
– second printing edition (2001).

[1997] *Suiji Guocheng Xindaoyin (A New Introduction to Stochastic Processes)* (in Chinese, with Rong Wu), World Scientific, Singapore (1997), 218 pp.

[2001] *Introduction to Random Time and Quantum Randomness* (with Jean-Claude Zambrini), Monographs of the Portugese Math. Soc., McGraw-Hill, Lisbon (2001), viii+177 pp.
– new edition, World Scientific, River Edge, NJ (2003), xii+211 pp.

[2004] *Chance and Choice. Memorabilia.* World Scientific, Hackensack, NJ, 2004, x+316 pp.

Other Publications

[1954] B. V. Gnedenko and A. N. Kolmogorov, *Limit Distributions for Sums of Independent Random Variables*, translated from the Russian and annotated by K. L. Chung, with appendix by J. L. Doob, Addison-Wesley Publishing Company, Inc., Cambridge, Mass. (1954), ix+264 pp.
– revised edition, translated from Russian, annotated and revised by K. L. Chung with an unpublished manuscript by Pao Lu Hsu and appendices by J. L. Doob, (1968), ix+293 pp.

[1982] *P. L. Hsu Collected Papers*, edited and with a preface by K. L. Chung, with the cooperation of Ching-Shui Cheng and Tse-Pei Chiang, Springer-Verlag (1983), xii+589 pp.

[1981–90, 1991–92] *Proceedings of the Seminar on Stochastic Processes,* edited with E. Çinlar and R. K. Getoor: 1981, 1982, 1983, 1984, 1985, 1986, 1987, 1988, 1989; edited with E. Çinlar and M. J. Sharpe: 1991, 1992, Managing editors: J. Glover (1985, 1986, 1987, 1988), P. J. Fitzsimmons and R. J. Williams (1989), P. Fitzsimmons, S. Port and T. Liggett (1991), R. F. Bass and K. Burdzy (1992), Birkhäuser, Boston.

PART 2
Comments on Selected Works

Sums of independent random variables and Markov chains by Naresh Jain

Professor Kai Lai Chung's contributions to probability theory have had a major influence on several areas of research in the subject. I will restrict my comments to some of his work in two areas, sums of independent random variables and the theory of Markov chains, which led to a significant amount of further work, including some of my own.

Sums of Independent Random Variables

Kai Lai has made many outstanding contributions to this field, but I would like to concentrate on his 1948 paper [2]. If $X_1, X_2, \ldots$ is a sequence of real-valued independent random variables, and $S_n = X_1 + X_2 + \cdots + X_n$, $n \geq 1$, denotes the sequence of partial sums, then the almost sure behavior of "large values" of $\{S_n\}$ was very well understood. Indeed, in the independent and identically distributed (i.i.d) case, Hartman and Wintner in 1941 [11] had already proved their celebrated law of the iterated logarithm: $EX_1 = 0$, $EX_1^2 = 1$, imply

$$\limsup_{n\to\infty} \frac{S_n}{(2n \log\log n)^{1/2}} = 1, \ a.s. \tag{1}$$

In a more general non-i.d. context, Feller in 1943 [8] had written almost the final word. Chung [2] observed that if S_n is replaced by $|S_n|$ in (1), the assertion remains valid. However, if S_n^* denotes $\max_{1\leq j\leq n} |S_j|$, then the behavior of "small values" of S_n^* had yet to be understood. He studied this problem in [2] including the non-i.d. situation and proved that if $EX_j \equiv 0$, and $E(S_n^2)$ is denoted by s_n^2, then under a natural third moment assumption

$$\liminf_{n\to\infty} \frac{S_n^*}{s_n(\log\log s_n)^{-1/2}} = 8^{-1/2}\pi, \ a.s. \tag{2}$$

To prove this result, he obtained the very profound probability estimate

$$P(S_n^* < cs_n) = \frac{4}{\pi}\sum_{j=0}^{\infty} \frac{(-1)^j}{2j+1} \exp\left(-\frac{(2j+1)^2\pi^2}{8c^2}\right)$$
$$+O((\log\log s_n/\log s_n)^{1/2}). \tag{3}$$

This contains the probability distribution for a standard one-dimensional Brownian motion process $\{B_t, t \geq 0\}$ if S_n^*/s_n is replaced by $\max_{0\leq t\leq 1}|B_t|$ on the left-side and the second term is replaced by zero on the right-side.

In the i.i.d. case, two questions arose after Chung's work. The first one was raised by Chung himself: If $EX_1 = 0$, $EX_1^2 = 1$, does (2) hold without any further assumptions with $s_n^2 = n$? The second natural question was to obtain the analogue of (2) if X_1 is in the domain of attraction of a stable law.

As to the first question, several papers appeared on the subject getting close to the conditions stipulated by Chung. The question was finally settled in the affirmative by Jain and Pruitt in 1975 [16]. The probability distribution given in (3) by Chung played a key role in the final solution. For the second question, if the index of stability $\alpha < 2$, it was not even clear if one should expect an analogue of (2).

Fristedt had already observed that an analogue of (1) could not exist if $\alpha < 2$. However, Fristedt and Pruitt [9] and Jain and Pruitt [15] showed, under different conditions on X_1, the existence of a real sequence $\{b_n\}$ increasing to infinity such that $\liminf(S_n^*/b_n) = c$, a.s., with $0 < c < \infty$, but it was not clear if the constant c depended on the distribution of X_1 or on the limit distribution alone.

Donsker and Varadhan [6] approached these problems through their large deviations probability estimates for stable processes and obtained explicit expressions for the limit constants. Jain [13] was then able to show through an invariance principle that the limit constant for the liminf behavior of (S_n^*/b_n) is the same as for the relevant stable process obtained by Donsker and Varadhan [6].

The story by no means ends here. For a two-parameter Brownian motion $B(s,t)$, $0 \leq s,t \leq 1$, the leading term of the "small ball" probability estimate for $P(\max_{0\leq s,t\leq 1}|B(s,t)| \leq c)$, as $c \downarrow 0$, is of great interest and turned out to be a challenging problem, which was solved by Bass [1] and Talagrand [21]. Much work has been done by other authors for parameter dimension larger than 2; these results, however, are not as definitive as in the two-dimensional parameter case. Many difficult questions still remain to be answered and we can expect these investigations to continue, owing their origins to Chung's pioneering work.

Markov Chains

It is difficult to imagine that anybody working in the area of Markov processes would not be familiar with Chung's monograph: Markov Chains with Stationary Transition Probabilities [3]. This monograph deals with countable state Markov chains in both discrete time (Part I) and continuous time (Part II). Much of Kai Lai's fundamental work in the field is included in this monograph. My comments will be confined to Part I. Here, for the first time, Kai Lai gave a systematic exposition of the subject which includes classification of states, ratio ergodic theorems, and limit theorems for functionals of the chain.

For a general state space, Doeblin had given a classification scheme in a seminal paper in 1937 [5]. This and other work of Doeblin had a major impact on the field and led to further developments by Chung [4], Doob [7], Harris [10], and Orey [18], [19]. In the early sixties there were a number of basic ingredients of a general state space theory that would lead to an exact counterpart of Part I of Chung [3]. Much fundamental ground work, including positive recurrence (the so-called Doeblin's condition), was done by Doob [7]. Harris [10] introduced his recurrence condition: There exists a nonzero σ-finite measure φ on the state space S such that $\varphi(E) > 0$ implies that starting from every $x \in S$, E is visited infinitely often a.s. He proved the existence of a (unique) σ-finite invariant measure π under this condition. If $\pi(S) = +\infty$, one could call the process null-recurrent, and one could ask if $\pi(E) < \infty$ implied $P^n(x, E) \to 0$ for every $x \in S$, as $n \to \infty$; here $P^n(x, E)$ denotes the n-step transition probability from x to E. This result was conjectured by Orey [18] and was a natural extension to the general state space situation of the corresponding well-known result for a countable state chain. Under Harris's recurrence condition one could also ask if an analogous ratio ergodic theorem was true, namely, if $\pi(F) > 0$ and $\pi(E) < \infty$, does

$$\sum_{j=1}^{n} P^j(x, E) / \sum_{j=1}^{n} P^j(y, F) \to \pi(E)/\pi(F) \tag{4}$$

as $n \to \infty$ for *all* $x, y \in S$? These questions were answered in [12] under Chung's guidance. Chung gave an example, reported in [12], to show that in the general case, as opposed to the countable state space case, (4) is true only for π-almost all x, y, and not for *all* x, y. The situation is different when the state space is not countable because one could stay in a π-null set for a rather long time! A little later, Jain and Jamison [14] introduced an irreducibility condition: There exists a nonzero σ-finite measure φ on S such that $\varphi(E) > 0$ implies that starting from every $x \in S$, the process visits E with positive probability. In this paper they essentially brought the program of Doeblin [5] and Chung [4] to completion. Chung's influence can be seen throughout these works. For other work in the area one can refer to monographs by Orey [19], Revuz [20] and Meyn and Tweedie [17].

References

[**1**] Bass, R. (1988). Probability estimates for multiparameter Brownian processes, Ann. Probab. **16**, 251-264.

[**2**] Chung, K.L. (1948). On the maximum partial sums of independent random variables, Trans. Amer. Math. Soc. **64**, 205-233.

[**3**] Chung, K.L. (1960). *Markov chains with stationary transition probabilities*, Grund. Math. Wiss. **104**, Springer-Verlag, New York/Berlin. Second Edition (1967).

[4] Chung, K.L. (1964). The general theory of Markov processes according to Doeblin, Z. Wahr. Verw. Gebiete **2**, 230-254.

[5] Doeblin, W. (1937). Sur les propriétés asymptotiques de mouvement régis par certain type de chaines simples. Bull. Math. Soc. Roum. Sci. **39**, No. 1, 57-115; No. 2, 3-61.

[6] Donsker, M.D. and Varadhan, S.R.S. (1977). On laws of the iterated logarithm for local times, Comm. Pure Appl. Math. **30**, 707-753.

[7] Doob, J.L. (1953). *Stochastic Processes*, Wiley and Sons, New York.

[8] Feller, W. (1943). The general form of the so-called law of the iterated logarithm, Trans. Amer. Math. Soc. **54**, 373-402.

[9] Fristedt, F. and Pruitt W.E. (1971). Lower functions for increasing random walks and subordinators, Z. Wahr. Verw. Gebiete **18**, 167-182.

[10] Harris, T.E. (1955). The existence of stationary measures for certain Markov processes, Proc. Third Berkeley Symp. Math. Statist. Probab. Vol II, 113-124.

[11] Hartman, P. and Wintner, A. (1941). On the law of the iterated logarithm, Amer. J. Math. **63**, 169-176.

[12] Jain, N.C. (1966). Some limit theorems for a general Markov process, Z. Wahr. Verw. Gebiete **6**, 206-223.

[13] Jain, N.C. (1982). A Donsker-Varadhan type of invariance principle, Z. Wahr. Verw. Gebiete **59**, 117-138.

[14] Jain, N.C. and Jamison, B. (1967). Contributions to Doeblin's theory of Markov processes, Z. Wahr. Verw. Gebiete **8**, 19-40.

[15] Jain, N.C. and Pruitt, W.E. (1973). Maxima of partial sums of independent random variables, Z. Wahr. Verw. Gebiete **27**, 141-151.

[16] Jain, N.C. and Pruitt, W.E. (1975). The other law of the iterated logarithm, Ann. Probab. **3**, 1046-1049.

[17] Meyn, S.P. and Tweedie, R.L. (1993). *Markov chains and stochastic stability*, Springer-Verlag, New York.

[18] Orey, S. (1959). Recurrent Markov chains, Pacific J. Math. **9**, 805-827.

[19] Orey, S. (1971). *Limit Theorems for Markov Chain Transition Probabilities*, Van Nostrand, New York.

[20] Revuz, D. (1975). *Markov Chains*, North-Holland, New York. Revised Edition (1984).

[21] Talagrand, M. (1994). The small ball problem for the Brownian sheet, Ann. Probab. **22**, 1331-1354.

Naresh Jain
School of Mathematics
University of Minnesota

Excursions, moderate Markov processes and probabilistic potential theory by Ronald Getoor

I first met Kai Lai Chung in 1955 when he gave a seminar talk at Princeton where I was an instructor at the time. I believe that he spoke about his work on Markov chains. After so many years I remember very little about the talk, but I clearly remember how impressed I was with the enthusiasm and energy displayed by the speaker. I had the pleasure of spending the academic year 1964–65 at Stanford and during that time Kai Lai and I became good friends. We had the opportunity to discuss mathematics in some depth and we often had lunch together. He had become interested in potential theory and had invited Marcel Brelot to visit Stanford during the spring quarter of 1965 and give a course on classical potential theory. By the end of the term he and I were the only ones still attending Brelot's lectures! During the 1970's we had an extensive correspondence about Markov processes, probabilistic potential theory and related topics. Interacting with Kai Lai on any level was always extremely stimulating and rewarding.

In what follows I am going to to comment on some of his work that was especially important and influential in areas that are of interest to me.

Excursions

During the early 1970's there was a considerable body of work on what might be called the general theory of excursions of a Markov process. Perhaps the definitive work in this direction was the paper of Maisonneuve [M]. Shortly thereafter Chung's paper [1976c] on Brownian excursions appeared. Some of his results had been announced earlier in [1975a]. Chung did not make use of the general theory; rather working by hand he made a deep study of the excursions of Brownian motion from the origin using the special properties of Brownian motion. This paper was a tour-de-force of direct methods for penetrating the mysteries of these excursions. Guided somewhat, it seems, by analogy with his previous work on Markov chains and inspired by Lévy's work he obtained a wealth of explicit formulas for the distributions of various random variables and processes derived from an excursion. I shall describe briefly a few of his results without reproducing the detailed explicit

expressions in the paper.

Let $B = B(t)$ denote one dimensional Brownian motion starting from the origin and let $Y = |B|$. Fix $t > 0$. Following Chung define

$$\gamma(t) = \sup\{s \leq t : Y(s) = 0\}; \beta(t) = \inf\{s \geq t : Y(s) = 0\}.$$

Then the intervals $(\gamma(t), \beta(t))$ and $(\gamma(t), t)$ are called the excursion interval straddling t and the interval of meandering ending at t respectively. Let $L(t) = \beta(t) - \gamma(t)$ and $L^-(t) = t - \gamma(t)$ denote the lengths of these intervals. Chung begins by giving a direct derivation of a number of results, originally due to Lévy, which lead to the joint distribution of $(\gamma(t), Y(t), \beta(t))$. Moreover based on his earlier work on Markov chains he is able to write these formulas in a particularly illuminating form. Define

$$Z^-(u) = Y(\gamma(t) + u), 0 \leq u \leq L^-(t); Z(u) = Y(\gamma(t) + u), 0 \leq u \leq L(t).$$

Then Z^- is called the meandering process and Z the excursion process. Theorem 4 gives the joint law of $\gamma(t)$ and Z^- while Theorem 6 contains the joint law of $\gamma(t), L(t)$ and Z. (Chung denotes both the meander process and the excursion process by Z; I have changed the notation for this exposition.) Chung then applies these results to calculate the distributions of various functionals of these processes. Particularly interesting is Theorem 7 which contains an explicit formula for the maximum of Z conditioned on $\gamma(t)$ and $L(t)$. A consequence is that

$$F(x) = 1 + 2\sum_{n=1}^{\infty}(1 - 2nx)e^{-n^2 x}, 0 < x < \infty$$

defines a distribution function! This is discussed in some detail. Other functionals were also studied. Of special interest to me is the occupation time of an interval $(a.b)$ during an excursion defined by

$$S(t, a.b) = \int_{\gamma(t)}^{\beta(t)} 1_{(a.b)}(Z(s))ds.$$

Among other things Chung showed that $S(t, 0, \varepsilon)/\varepsilon^2$ has a limiting distribution as $\varepsilon \downarrow 0$ and computed its first four moments. In [GS] it was shown that this distribution was the convolution of the first passage distribution $P(R \in ds)$ with itself where $R = \inf\{s : Y(s) = 1\}$.

References

[GS] R. K. Getoor and M. J. Sharpe, Excursions of Brownian motion and Bessel processes, *Z. Warsch. Verw. Gebiete,* **47** (1979), 83-106.

[M] B. Maisonneuve, Exit systems, *Ann. Prob.,* **3** (1975), 399-411.

Moderate Markov Processes

In the paper [1979a] some of the basic properties of a left continuous moderate Markov process were formulated and proved. It was more or less ignored when it first appeared, even though the importance of this class of processes was evident from the fundamental paper of Chung and Walsh [1969a] on time reversal of Markov processes. In the [1969a] paper it was called the moderately strong Markov property and the process was right continuous. To the best of my knowledge the terminology moderate Markov property first appeared in [1972c]. In 1987 Fitzsimmons [F] was able to modify somewhat the Chung-Walsh methods and so to construct a left continuous moderate Markov dual process for any Borel right process and given excessive measure, m, as duality measure. The Chung-Walsh theory corresponds to m being the potential of a measure μ which served as a fixed initial distribution. More importantly Fitzsimmons showed that this dual was a powerful tool in studying the potential theory of the underlying Borel right process. Consequently there was renewed interest in left continuous moderate Markov processes and the Chung-Glover paper [1979a] was immediately relevant. It has become the basic reference for properties of these processes.

Reference

[F] P. J. Fitzsimmons, Homogeneous random measures and a weak order for excessive measures of a Markov process, *Trans. Amer. Math. Soc.*, **303** (1987), 431-478.

Probabilistic Potential Theory

The paper [1973a] was perhaps Chung's most influential contribution to what is commonly known as probabilistic potential theory. (This excludes his work on gauge theorems and Schrödinger equations). In it he obtained a beautiful expression for the equilibriuim distribution of a set in terms of the last exit distribution from the set and the potential kernel density of the underlying process. He emphasized and clearly stated that his approach involved working directly with the last exit time. This was an important innovation since such times are not stopping times and so were not part of the available machinery at that epoch. Immediately following Chung's paper (more precisely its announcement) and inspired by it, Meyer [M73] and Getoor and Sharpe [GS73] obtained, similar results under different hypotheses. Numerous authors then developed techniques for handling last exit and more general times which became part of the standard machinery of Markov processes. In two additional papers [1975c] and (with K. Murali Rao) [1980a], conditions were given under which the equilibrium measure obtained from the last exit distribution is a

multiple of the measure of minimum energy as in classical situations. In [1980a] symmetry was not assumed and so a modified form of energy was introduced in order to obtain reasonable results. Additional implications in potential theory of the hypotheses he had introduced in [1973a] and also their relationship with the more common duality hypotheses were explored with K. Murali Rao in [1980f] and with Ming Liao and Rao in [1984c]. Of particular importance was the result giving sufficient conditions for the validity of Hunt's hypothesis (B) in [1980f]. These four papers were very original, but for some reason they were not as influential as the paper [1973a].

For historical reasons I should point out that the relationship between the equilibrium measure and the last exit distribution had appeared a few years earlier in Port and Stone's memoir on infinitely divisible processes – see sections 8 and 11 of [PS71]. One may wonder why Chung's paper [1973c], was immediately so influential while the result in Port and Stone was hardly noticed at the time. Certainly it was unknown to Chung and evidently Meyer was also unaware of it. The most likely reasons for this are two fold: (1) The result in Port and Stone was buried in a memoir of just over two hundred pages; in addition their proof of the integral condition for the transience or recurrence of an infinitely divisible process attracted the most attention at the time. (2) In Chung it was the main result of the paper, it was clearly stated and proved by a direct easily understood argument.

I shall now explain in a bit more detail what Chung did. I'll try to emphasize the ideas leaving aside technicalities. So suppose that $X = (X_t, P^x)$ is a Hunt process taking values in a locally compact, separable Hausdorff space E. If $B \in \mathcal{E}$, the σ-algebra of Borel subsets of E, define the hitting time, T_B, and the last exit time, λ_B, of B by

$$T_B = \inf\{t > 0 : X_t \in B\}, \quad \lambda_B = \sup\{t > 0 : X_t \in B\}$$

where the inf (resp. sup) of the empty set is ∞ (resp. 0). Let $U(x,B) = E^x \int_0^\infty 1_B(X_t)dt$ denote the potential kernel of X and suppose that $U(\cdot, K)$ is bounded for K compact; in particular X is transient. For the moment suppose X is Brownian motion (BM) in $\mathbb{R}^d, d \geq 3$. Then $U(x,B) = \int_B u(x,y)dy$ where $u(x,y) = c_d|x-y|^{2-d}$ is the Newtonian potential kernel appropriately normalized. A classical result in potential theory states that if $K \subset \mathbb{R}^d$ is compact and has positive (Newtonian) capacity, then there exists a unique measure μ_K, called the equilibrium measure or distribution of K, carried by K and whose potential.

$$p_K(x) = U\mu_K(x) = \int u(x,y)\mu_k(dy) \tag{1}$$

is less than or equal to 1 everywhere and takes the value 1 on K. Actually $p_K \equiv 1$ on K only if K is regular; in general there may be an exceptional subset of K of capacity zero on which $p_K < 1$. The function p_K is called the equilibrium potential of K and may be characterized as the unique superharmonic function v on $\mathbb{R}^d$ such that $0 \leq v \leq 1, v$ is harmonic on $\mathbb{R}^d \backslash K$ and $\{v < 1\} \cap K$ has capacity zero – $v \equiv 1$

on K if K is regular. Evidently Kakutani [K44] was the first person to note that

$$p_K(x) = P^x(T_K < \infty) = P^x(X_t \in K \text{ for some } t > 0). \tag{2}$$

Then one may ask for what class of Borel sets $B \subset \mathbb{R}^d$ does there exist a measure μ_B such that

$$P^x(T_B < \infty) = \int u(x,y)\mu_B(dy) \tag{3}$$

and what can be said about μ_B. This is the equilibrium problem as stated in the first paragraph of Chung's paper.

Now return to the situation in which X is a Hunt process as described in the first few sentences of the preceding paragraph. For $B \in \mathcal{E}$ recall the definitions of the hitting time T_B and the last exit time λ_B. The set B is transient provided $P^x(\lambda_B < \infty) = 1$ for all x. Also note that

$$p_B(x) = P^x(T_B < \infty) = P^x(\lambda_B > 0).$$

Fix B transient and let $p = p_B$. It is easily checked that p is excessive and $P_t p \to 0$ as $t \to \infty$. Here $P_t = (P_t(x, \cdot))$ is the transition semigroup of X. Formally, from semigroup theory $\frac{1}{\epsilon}(p - P_\epsilon p) \to -\mathcal{G}p$ where $\mathcal{G}$ is the "generator" of (P_t) and $p = U(-\mathcal{G}p)$. Here U is the potential kernel of X as defined above. Of course, in general p is not in the domain of $\mathcal{G}$. However, if we want to represent p as the potential of something, then one expects it to be some sort of limit of $p_\epsilon = \frac{1}{\epsilon}(p - P_\epsilon p)$ as $\epsilon \downarrow 0$. This idea had been used by McKean and Tanaka [MT61], Volkonski [V60] and Šur [S61] to represent excessive functions as potentials of additive functionals. More relevant to the present discussion, using the same basic idea, Hunt [H58] had shown for what are now called Hunt processes satisfying, in addition, the existence of a nice dual process and subject to a type of Feller condition and a transience hypothesis, that if B has compact closure, then (3) holds where now $u(x,y)$ is the potential density associated with X and its dual, in particular $U(x,dy) = u(x,y)m(dy)$ where m is the duality measure – Lebesgue measure when X is Brownian motion.

Chung's key observation was to note that $p - P_\epsilon p = P^\cdot(\lambda_B > 0) - P^\cdot(\lambda_B > \epsilon) = P^\cdot(0 < \lambda_B \leq \epsilon)$. Suppose $f \geq 0$ is a bounded continuous function and for simplicity write $\lambda = \lambda_B$. Then by the Markov property

$$\begin{aligned} U[f(p - P_\epsilon p)] &= E^\cdot \int_0^\infty f(X_t) P^{X(t)}(0 < \lambda \leq \epsilon)\,dt \\ &= E^\cdot \int_0^\infty f(X_t) 1_{\{0 < \lambda\circ\theta_t \leq \epsilon\}}\,dt. \end{aligned}$$

Here θ_t is the shift operator which shifts the origin of the path from 0 to t so that $X_s \circ \theta_t = X_{s+t}$, $s \geq 0$. It is easily checked that $\lambda \circ \theta_t = (\lambda - t)^+$. Plugging this into the last integral and recalling that $p_\epsilon = \frac{1}{\epsilon}(p - P_\epsilon p)$ one finds

$$\begin{aligned} U[fp_\epsilon] &= \frac{1}{\epsilon}\, E^\cdot\Big[\int_{(\lambda-\epsilon)^+}^{\lambda} f(X_t)\,dt; \lambda > 0\Big] \\ &\to E^x[f(X_{\lambda-}), \lambda > 0] \text{ as } \epsilon \downarrow 0. \end{aligned} \tag{4}$$

Now suppose that there exists a Radon measure m on E such that $U(x,dy) = u(x,y)m(dy)$. Then Chung imposed analytic conditions on the potential density $u(x,y)$ which implied the existence of a measure μ_B such that

$$\begin{aligned} U[fp_\epsilon](x) &= \int u(x,y)f(y)p_\epsilon(y)m(dy) \\ &\to \int u(x,y)f(y)\mu_B(dy) = U[f\mu_B](x) \text{ as } \epsilon \downarrow 0 \end{aligned}$$

for all bounded continuous f with compact support. Combining this with (4) we obtain

$$E^x[f(X_{\lambda-}); \lambda > 0] = U[f\mu_B](x), \tag{5}$$

and taking a sequence of such f increasing to **1**,

$$p_B(x) = P^x[T_B < \infty] = P^x[\lambda_B > 0] = U\mu_B(x). \tag{6}$$

Defining the last exit distribution $L_B(x,dy) = P^x[X_{\lambda-} \in dy, \lambda > 0]$, (5) implies that

$$L_B(x,dy) = u(x,y)\mu_B(dy). \tag{7}$$

This formula (7) is the celebrated result of Chung which gives the probabilistic meaning of the equilibrium measure μ_B. The measure μ_B is carried by $\overline{B}$, even by ∂B when X has continuous paths. Under Chung's or Hunt's hypotheses μ_B is a Radon measure; more generally, under duality without Feller conditions μ_B is σ-finite.

Let me derive a simple consequence of (5) and for simplicity I shall suppose X is Brownian motion in $\mathbb{R}^d, d \geq 3$. Let $B \subset \mathbb{R}^d$ be transient, for example $\overline{B}$ compact. As before $\lambda = \lambda_B$. Since the paths are continuous, (5) and the Markov property imply that

$$E^x[f(X_\lambda); 0 < \lambda \leq t] = Uf\mu_B(x) - P_t Uf\mu_B(x)$$

for $t > 0$ and f bounded with compact support. Now $P_t(x,dy) = g_t(y-x)dy$ where g_t is the familar Gauss kernel. Hence

$$E^x[f(X_\lambda); 0 < \lambda \leq t] = \int \int_0^t ds\ g_s(y-x)f(y)\mu_B(dy).$$

Integrating over $\mathbb{R}^d$ we obtain since g_s is a probability density

$$\int_{\mathbb{R}^d} dx E^x[f(X_\lambda); 0 < \lambda \leq t] = t \int f\ d\mu_B;$$

that is,

$$P^m[X_\lambda \in dy, \lambda \in dt] = dt\mu_B(dy), t > 0 \tag{8}$$

where m is Lebesgue measure. Thus X_λ and λ are independent under the σ-finite measure P^m and their joint distribution under P^m is the product of μ_B and Lebesgue measure. To my mind, this is one of the nicest probabilistic interpretations

of the equilibrium measure for Brownian motion. Actually this is valid in much more generality. For example, if X has a strong dual and the duality measure m is invariant, then

$$P^m(X_{\lambda-} \in dy, \lambda \in dt) = dt\mu_B(dy), t > 0. \tag{9}$$

See [GS73]. In particular this holds for transient Lévy processes in $\mathbb{R}^d$ whose potential kernel is absolutely continuous. In general if m is not invariant, then X_λ and λ are not independent under P^m.

References

[**GS73**] R. K. Getoor and M. J. Sharpe, Last exit times and additive functionals, *Ann. Prob.* **1** (1973), 550-569.

[**H58**] G. A. Hunt, Markoff processes and potentials III, *Illinois J. Math.* **2** (1958), 151-213.

[**K44**] S. Kakutani, Two dimensional Brownian motion and harmonic functions, *Proc. Imp. Acad. Tokyo,* **20** (1944), 706-714.

[**MT61**] H. P. McKean, Jr. and H. Tanaka, Additive functionals of the Brownian path, *Memoirs Coll. Sci., Univ. Kyoto, Ser A,* **33** (1961), 479-506.

[**M73**] P. A. Meyer, Note sur l'interprétation des mesures d'équilbre, *Sém. de Prob. VII, Springer Lecture Notes Math.,* **321** (1973), 210-216.

[**PS71**] S. C. Port and C. J. Stone, Infinitely divisible processes and their potential theory I., *Ann. Inst. Fourier (Grenoble),* **21** (1971), 157-275.

[**S61**] M. G. Šur, Continuous additive functionals of a Markov process (in Russian), *Dokl. Akad. Nauk. SSSR,* **137** (1961), 800-803. English translation, *Soviet Math.* **2** (1961), 365-368.

[**V60**] V. A. Volkonskii, Additive functionals of Markov processes (in Russian), *Tr. Mosk. Mat. Obšč.* **9** (1960), 143-189. English translation, *Selected translations in Math. Stat. and Prob.,* AMS, **5** (1965), 127-178.

Ronald K. Getoor
Department of Mathematics
University of California, San Diego

This page intentionally left blank

Stopped Feynman-Kac functionals and the Schrödinger equation by Ruth Williams

In the late 1970s, Kai Lai Chung began investigating connections between probability and the (reduced) Schrödinger equation:

$$\frac{1}{2}\Delta u(x) + q(x)u(x) = 0 \quad \text{for } x \in \mathbb{R}^d, \tag{1}$$

where q is a real-valued, Borel measurable function on $\mathbb{R}^d$ and Δ is the d-dimensional Laplacian. His work in this area extended over the next 15 years or so. It included collaborations with several colleagues and students, and inspired the work of others. His book [CZ], "From Brownian Motion to Schrödinger's Equation", written with Zhongxin Zhao, is a compilation and refinement of much of the research conducted in this area up through 1994.

In the following, I will describe some of the background and early advances in this research involving connections with Brownian motion. A complementary article written by Michael Cranston, which also appears in this volume, focuses on related developments involving connections with conditioned Brownian motion. My account is not meant to be exhaustive, but rather to provide a sample of some of the intriguing aspects of the topic and to illustrate the pivotal role that Kai Lai Chung played in some of the developments. My description is necessarily influenced by my own personal recollections.

Background

Stimulated by Feynman's [Fe] proposed "path integral" solution of the complex time-dependent Schrödinger equation, for a Borel measurable function $q : \mathbb{R} \to \mathbb{R}$ satisfying $q \leq 0$, Kac [Ka, Kb] considered the following multiplicative functional of one-dimensional Brownian motion B:

$$e_q(t) = \exp\left(\int_0^t q(B_s)ds\right), \quad t \geq 0, \tag{2}$$

This functional can also be defined for suitable Borel measurable functions $q : \mathbb{R}^d \to \mathbb{R}$ and B a d-dimensional Brownian motion or even a d-dimensional diffusion process. Such functionals are now called *Feynman-Kac functionals.*

Consider a continuous, bounded function $q : \mathbb{R}^d \to \mathbb{R}$ and a continuous, bounded function $g : \mathbb{R}^d \to \mathbb{R}$. If $\psi : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}$ is a continuous, bounded function, with continuous partial derivatives $\frac{\partial \psi}{\partial t}$, $\frac{\partial \psi}{\partial x_i}$, and $\frac{\partial^2 \psi}{\partial x_i \partial x_j}$, $i, j = 1, \ldots, d$, defined on $(0,\infty) \times \mathbb{R}^d$, such that the following time-dependent Schrödinger equation holds:

$$\frac{\partial \psi(t,x)}{\partial t} = \frac{1}{2}\Delta\psi(t,x) + q(x)\psi(t,x), \quad t > 0, \ x \in \mathbb{R}^d, \tag{3}$$

with initial condition $\psi(0,x) = g(x)$, $x \in \mathbb{R}^d$, and where

$$\Delta\psi(t,x) = \sum_{i=1}^{d} \frac{\partial^2 \psi}{\partial x_i^2}(t,x), \tag{4}$$

then it can be shown (for example by using Itô's formula), that

$$\psi(t,x) = E^x\left[e_q(t) g(B_t)\right], \quad t \geq 0, \ x \in \mathbb{R}^d. \tag{5}$$

Here E^x denotes the expectation operator under which B is a d-dimensional Brownian motion starting from x.

Kac [Ka, Kb] was interested in (3) when $d = 1$ and $q \leq 0$. However, rather than considering this equation directly, he worked with a reduced Schrödinger equation similar to (1) (with $q - s$ in place of q and $d = 1$), obtained by formally taking the Laplace transform (with parameter s) in equation (3) to eliminate the time variable t. Under mild conditions, for example q is bounded and continuous in addition to being non-positive, Kac [Kb] showed that the Laplace transform of the right member of (5) with $g = 1$ satisfies this reduced Schrödinger equation.

In the late 1950's and early 1960's, in developing a potential theory for Markov processes, Dynkin (cf. [D], Chapter XIII, §4, Theorem 13.16), and others, considered expressions of the form

$$E^x\left[\exp\left(\int_0^\tau q(X_s)ds\right) f(X_\tau)\right], \quad x \in \overline{D}, \tag{6}$$

where X is a diffusion process in $\mathbb{R}^d$, $\tau = \inf\{s > 0 : X_s \notin D\}$ is the first exit time of X from a bounded domain D in $\mathbb{R}^d$, $\overline{D}$ is the closure of D, f is a continuous function defined on the boundary of D, and $q \leq 0$ is Hölder continuous and bounded on D. Here P^x and E^x denote probability and expectation operators, respectively, for X starting from $x \in \overline{D}$. The domain D is *regular* if

$$P^x(\tau = 0) = 1 \quad \text{for each } x \in \partial D.$$

Under suitable assumptions on X and assuming that D is regular, Dynkin showed that expressions of the form (6) yield continuous functions on $\overline{D}$ that satisfy the equation

$$\mathcal{L}u(x) + q(x)u(x) = 0 \quad \text{for } x \in D, \tag{7}$$

with continuous boundary values given by f, where $\mathcal{L}$ is the infinitesimal generator of the diffusion process X. The assumption that $q \leq 0$ implies that the action of the stopped Feynman-Kac functional,

$$e_q(\tau) = \exp\left(\int_0^\tau q(X_s)ds\right), \tag{8}$$

is to "kill" the diffusion process at a state dependent exponential rate given by $-q$ up until the stopping time τ. The assumption that q is non-positive ensures that the mean value of the stopped Feynman-Kac functional is always finite; in fact, it is bounded by one. In contrast, Khasminskii [Ks] considered the case where q is *non-negative*. In this case, the action of the stopped Feynman-Kac functional can be interpreted as "creating mass" at a state dependent exponential rate given by q up until the stopping time τ. Accordingly, the expression in (6) can fail to be finite if the domain is sufficiently large. Indeed, the results of Khasminskii [Ks] imply that, under similar conditions to those imposed by Dynkin except that $q \geq 0$, the expression in (6) is finite for all $x \in \overline{D}$ if and only if there is a continuous function u that is strictly positive on $\overline{D}$ that satisfies (7).

It was not until the work of Chung [Ca] that probabilistic solutions of (1) for general (signed) q became an object of considerable interest. The question of how the oscillations of such a q affect the behavior of the stopped Feynman-Kac functional (8) is an intriguing one; in particular, killing of mass in some locations may cancel creation of mass in others. The next section describes some of K. L. Chung's investigations on stopped Feynman-Kac functionals with general q.

Feynman-Kac Gauge and Positive Solutions of the Schrödinger Equation

One-dimensional diffusions

Kai Lai Chung initiated his research on stopped Feynman-Kac functionals in [Ca] by considering a one-dimensional diffusion process (i.e., continuous strong Markov process) X and the functional (8) with bounded, Borel measurable $q : \mathbb{R} \to \mathbb{R}$ and $\tau = \tau_b \equiv \inf\{t > 0 : X_t = b\}$ for $b \in \mathbb{R}$. Assume that $P^x(\tau_b < \infty) = 1$ for each $x \in \mathbb{R}$ and $b \in \mathbb{R}$, and define

$$v(x,b) = E^x\left[\exp\left(\int_0^{\tau_b} q(X_s)\,ds\right)\right], \quad x \in \mathbb{R},\ b \in \mathbb{R}. \tag{9}$$

Two fundamental properties of v are that $0 < v(x,b) \leq \infty$ for all $x, b \in \mathbb{R}$, and for any $a < b < c$ or $c < b < a$ in $\mathbb{R}$,

$$v(a,b)v(b,c) = v(a,c). \tag{10}$$

The latter follows from the strong Markov property. The above properties lead to the following lemma.

Lemma 1. *For fixed $b \in \mathbb{R}$,*

(a) if $v(x,b) < \infty$ for some $x < b$, then $v(x,b) < \infty$ for all $x < b$,
(b) if $v(x,b) < \infty$ for some $x > b$, then $v(x,b) < \infty$ for all $x > b$.

Chung [Ca] introduced the following measures of finiteness of v:

$$\alpha = \inf\{b \in \mathbb{R} : v(x,b) < \infty \text{ for all } x > b\}, \tag{11}$$

$$\beta = \sup\{b \in \mathbb{R} : v(x,b) < \infty \text{ for all } x < b\}, \tag{12}$$

and showed the next two results.

Lemma 2.

$$\alpha = \sup\{b \in \mathbb{R} : v(x,b) = \infty \text{ for all } x > b\}, \tag{13}$$

$$\beta = \inf\{b \in \mathbb{R} : v(x,b) = \infty \text{ for all } x < b\}. \tag{14}$$

Furthermore, if $\beta \in \mathbb{R}$, then $v(x,\beta) = \infty$ for all $x < \beta$, and, if $\alpha \in \mathbb{R}$, then $v(x,\alpha) = \infty$ for all $x > \alpha$.

Theorem 1. *The following conditions (a)–(c) are equivalent.*

(a) $\beta = +\infty$.
(b) $\alpha = -\infty$.
(c) For all $a; b \in \mathbb{R}$,

$$v(a,b)v(b,a) \leq 1.$$

Following on from this, Chung and Varadhan [CV] proved the next two theorems for $q : \mathbb{R} \to \mathbb{R}$ bounded and continuous, and X a one-dimensional Brownian motion.

Theorem 2. *Fix $b \in \mathbb{R}$. Suppose that $v(x,b) < \infty$ for some, hence all, $x < b$. Then $u(x) = v(x,b)$ is twice continuously differentiable for $x \in (-\infty, b)$, continuous for $x \in (-\infty, b]$, and u satisfies the reduced Schrödinger equation:*

$$\frac{1}{2}u''(x) + qu(x) = 0, \quad x \in (-\infty, b), \tag{15}$$

with the boundary condition

$$u(b) = 1. \tag{16}$$

Theorem 3. *The following conditions (a)–(e) are equivalent.*

(a) There is a twice continuously differentiable, strictly positive function u satisfying (15) with $b = +\infty$.
(b) $\beta = +\infty$.
(c) $\alpha = -\infty$.
(d) For all $a, b \in \mathbb{R}$,

$$v(a,b)v(b,a) \leq 1. \tag{17}$$

(e) There is some pair of real numbers a, b, $a \neq b$, such that (17) holds.

The equivalence of (a), (b) and (c) is an analogue of Khasminskii's [Ks] results, but with $d = 1$, $D = (-\infty, \infty)$ and q being allowed to change sign. Further discussion of this one-dimensional case for Brownian motion can be found in Chapter 9 of the book by Chung and Zhao [CZ].

Multidimensional Brownian Motion

Chung soon moved on to consider multidimensional Brownian motions and domains of finite Lebesgue measure in the work [CR] with K. Murali Rao. This paper appeared in the proceedings of the first "Seminar on Stochastic Processes", held at Northwestern University in 1981. This series of annual conferences was initiated by K. L. Chung, E. Çinlar and R. K. Getoor. The "Seminars" have grown in size over the years, but the novel format of a few invited talks, with ample time reserved for less formal presentations and discussions, has persisted and is one of the attractions of these annual meetings held over two and a half days.

The paper [CR] was a significant advance. In particular, it contained the the first "gauge theorem". It is stated in its original form below and then some generalizations are mentioned.

For this, assume that B is a d-dimensional Brownian motion ($d \geq 1$), P^x and E^x are probability and expectation operators, respectively, for B starting from $x \in \mathbb{R}^d$, $q : \mathbb{R}^d \to \mathbb{R}$ is a bounded, Borel measurable function, D is a domain in $\mathbb{R}^d$ with closure $\overline{D}$ and boundary ∂D, m is d-dimensional Lebesgue measure, and $f : \partial D \to \mathbb{R}$ is a bounded, Borel measurable function satisfying $f \geq 0$. Let

$$\tau_D = \inf\{t > 0 : B_t \notin D\}, \tag{18}$$

the first exit time of B from D. Define

$$u(D, q, f; x) = E^x \left[\exp\left(\int_0^{\tau_D} q(B_s)ds \right) f(B_{\tau_D}); \tau_D < \infty \right], \quad x \in \overline{D}. \tag{19}$$

When $m(D) < \infty$, $P^x(\tau_D < \infty) = 1$ for all $x \in \overline{D}$ (cf. [CZ], Theorem 1.17), and the qualifier $\tau_D < \infty$ in (19) may be omitted. The following is Theorem 1.2 in Chung and Rao [CR].

Theorem 4. *Suppose the domain D satisfies $m(D) < \infty$. If $u(D, q, f; \cdot) \not\equiv \infty$ in D, then it is bounded in $\overline{D}$.*

While visiting Chung at Stanford University in the early 1980's, Zhongxin Zhao [Z] (see also [CZ], Theorems 5.19 and 5.20) extended this result by relaxing the assumptions on q and D. In particular, he showed that the conclusion of Theorem 4 continues to hold if the boundedness condition on q is relaxed to simply require that $q : \mathbb{R}^d \to \overline{\mathbb{R}}$ is a Borel measurable function satisfying

$$\lim_{\alpha \downarrow 0} \left[\sup_{x \in \mathbb{R}^d} \int_{|y-x| \leq \alpha} |G(x-y)q(y)|dy \right] = 0, \tag{20}$$

where $\overline{\mathbb{R}} = [-\infty, \infty]$ is the extended real line and for $x \in \mathbb{R}^d$,

$$G(x) = \begin{cases} |x|^{2-d} & \text{if } d \geq 3, \\ \ln \frac{1}{|x|} & \text{if } d = 2, \\ |x| & \text{if } d = 1. \end{cases} \tag{21}$$

The set of Borel measurable functions $q : \mathbb{R}^d \to \overline{\mathbb{R}}$ satisfying (20) is called the *Kato class* (on $\mathbb{R}^d$) and is usually denoted by K_d or J. Various properties of these functions, as well as analytic properties of associated weak solutions of the reduced Schrödinger equation (22), are described in an extensive paper of Aizenman and Simon [AS] which appeared shortly after the work [CR] of Chung and Rao. The paper [AS] also describes connections between the stopped Feynman-Kac functional and weak solutions of (22) under a spectral condition (cf. Theorem 6 below).

Neil Falkner [Fa] also visited Chung at Stanford in the early 1980's. During this time, Falkner proved a gauge theorem, when conditioned Brownian motion is used in place of Brownian motion, for bounded Borel measurable q and sufficiently smooth bounded domains D. (Zhao subsequently used conditioned Brownian motion in his work [Z].) For more details on the work in [Fa] and a discussion of subsequent generalizations, see the article by Michael Cranston in this volume and Chapter 7 of the book by Chung and Zhao [CZ].

The function $u(D, q, 1; \cdot)$ obtained by setting $f \equiv 1$ is called the *gauge* (function) for (D, q) and we say that (D, q) is *gaugeable* if this function is bounded on D.

Under the assumptions of Theorem 4 and assuming (D, q) is gaugeable, a second key result in the paper of Chung and Rao [CR] provides sufficient conditions for $u(D, q, f; \cdot)$ to be a twice continuously differentiable solution of the reduced Schrödinger equation in D:

$$\frac{1}{2}\Delta u(x) + q(x)u(x) = 0 \quad \text{for } x \in D, \tag{22}$$

with continuous boundary values given by f. As is usual in the theory of elliptic partial differential equations, to ensure two continuous derivatives for u, in dimensions two and higher, one imposes a stronger condition on q than simple continuity. For example, locally Hölder continuous functions are often used. For $d = 1$, let $C_1(D)$ denote the set of bounded, continuous functions $h : D \to \mathbb{R}$, and for $d \geq 2$, let $C_d(D)$ denote the set of bounded, continuous functions $h : D \to \mathbb{R}$ such that for each compact set $K \subset D$, there is are strictly positive, finite constants α, M such that

$$|h(x) - h(y)| \leq M|x - y|^\alpha, \quad \text{for all } x, y \in K. \tag{23}$$

The following theorem is proved in Chung and Rao [CR] for $d \geq 2$. For $d = 1$, they impose local Hölder continuity on q to obtain the result, but this condition can be relaxed by invoking a suitable analytic lemma for a Green potential, as was shown in Chung's book [Cb], Lectures from Markov Processes to Brownian Motion (see Proposition 4 of Section 4.7 of [Cb]). Note for this that for $d = 1$, $m(D) < \infty$ implies that D is a bounded interval.

Theorem 5. *Let D be a regular domain in $\mathbb{R}^d$ satisfying $m(D) < \infty$. Suppose that $q \in C_d(D)$ and $f : \partial D \to \mathbb{R}$ is bounded and continuous. Assume that (D, q) is gaugeable, i.e., $u(D, q, 1; \cdot) \not\equiv \infty$ in D. Then $u = u(D, q, f; \cdot)$ defined by (19) on $\overline{D}$ is twice continuously differentiable in D, continuous and bounded on $\overline{D}$, it satisfies*

(22) in D and $u = f$ on ∂D. Furthermore, u is the unique twice continuously differentiable solution of (22) that is continuous and bounded on $\overline{D}$ and agrees with f on the boundary ∂D.

This theorem has been generalized to situations where q is a Kato class function and (22) is interpreted in the weak sense of partial differential equation theory (cf. [CZ], Section 4.4)

Note that under the assumptions of the theorem above, if $f \geq 0$, then $u(D, q, f; \cdot)$ is a non-negative solution of (22), and if $f > 0$ on ∂D, then $u(D, q, f; \cdot) > 0$ on $\overline{D}$. One naturally expects there to be some relation between the existence of such positive solutions of (22) and the sign of

$$\lambda(D, q) = \sup_{\phi} \left[\int_D \left\{ -\frac{|\nabla\phi(x)|^2}{2} + q(x)\phi(x)^2 \right\} dx \right], \tag{24}$$

where the supremum is over all $\phi : D \to \mathbb{R}$ such that ϕ is infinitely continuously differentiable in D, has compact support in D and satisfies $\int_D \phi(x)^2 dx = 1$. The quantity $\lambda(D, q)$ is the supremum of the spectrum of the operator $\frac{1}{2}\Delta + q$ on $L^2(D)$ (cf. [CZ], Proposition 3.29). Indeed, there is a sharp relationship provided by the following theorem (see Theorem 4.19 of [CZ] for a proof).

Theorem 6. *Let D be a domain in $\mathbb{R}^d$ satisfying $m(D) < \infty$ and q be a Kato class function. Then (D, q) is gaugeable if and only if $\lambda(D, q) < 0$.*

For bounded domains D, in Theorem A.4.1 of [AS], Aizenman and Simon proved the "if" part of this theorem and that when $\lambda(D, q) < 0$, $u(D, q, f; \cdot)$ is a weak solution of (22), and its boundary values are given by f and they are assumed continuously if f is continuous and D is regular.

The work of Chung and Rao [CR] was the seed for much subsequent work on connections between the probabilistically defined quantity (19), gauges, and solutions of the reduced Schrödinger equation (22). Besides continuing his own work in this area, in the 1980's Chung had two students, Elton P. Hsu [CH, Ha, Hb] and Vassilus Papanicolaou [P], who worked on probabilistic representations for other boundary value problems associated with the reduced Schrödinger equation. A conjecture of Chung on equivalent conditions for finiteness of the gauge in terms of finiteness of $u(D, q, f; \cdot)$, when f is a non-negative function that is positive only on a suitable subset of the boundary, stimulated work of myself [W] (as a student of Chung) and then Neil Falkner [Fa] (as a visitor at Stanford). Falkner's work used conditioned Brownian motion which became an object of intense interest in its own right and for its connections with gauge theorems. For more on this fascinating subject, see the accompanying article by Michael Cranston. Other generalizations have also occurred, especially ones involving more general Markov processes than Brownian motion. The works related to this are too numerous to mention here.

Finally, on a personal note, I would like to thank Kai Lai Chung for the pleasure of our collaborations and for the many lively discussions I have enjoyed with him over the years.

References

[**AS**] M. Aizenman and B. Simon (1982). Brownian motion and Harnack's inequality for Schrodinger operators, *Comm. Pure Appl. Math.* **35** 209–273.

[**Ca**] K. L. Chung (1980). On stopped Feynman-Kac functional, *Séminaire de Probabilités XIV (1978/79), Lecture Notes in Mathematics,* **784,** Springer, New York, 347–356.

[**Cb**] K. L. Chung (1982). *Lectures from Markov processes to Brownian motion,* Springer-Verlag, New York.

[**CH**] K. L. Chung, and P. Hsu (1986). Gauge theorem for the Neumann problem, *Proceedings of the Seminar on Stochastic Processes, 1984 (Evanston, Ill., 1984),* Progr. Probab. Statist., **9,** Birkhäuser, Boston, MA, 63–70.

[**CR**] K. L. Chung and K. M. Rao (1981). Feynman-Kac functional and Schrodinger equation, *Seminar on Stochastic Processes,* Birkhäuser, Boston, MA, 1–29.

[**CV**] K. L. Chung and S. R. S. Varadhan (1980). Kac functional and Schrödinger equation, *Studia Mathematica,* **68** 249–260.

[**CZ**] K. L. Chung and Z. Zhao (1995). *From Brownian Motion to Schrödinger's Equation,* Springer-Verlag, New York.

[**D**] E. B. Dynkin (1965). *Markov Processes, Vol. II,* Springer-Verlag, New York.

[**Fa**] N. Falkner (1983). Feynman-Kac functional and positive solutions of $\frac{1}{2}\Delta u + qu = 0$, *Z. Wahrsch. verw. Gebiete,* **65** 19–34.

[**Fe**] R. P. Feynman (1948). Space-time approach to non-relativistic quantum mechanics, *Rev. Modern Phys.* **20** 367–387.

[**Ha**] P. Hsu (1985). Probabilistic approach to the Neumann problem, *Comm. Pure Appl. Math.* **38** 445–472.

[**Hb**] P. Hsu (1987). On the Poisson kernel for the Neumann problem of Schrödinger operators. *J. London Math. Soc.* **36** 370–384.

[**Ka**] M. Kac (1949). On distributions of certain Wiener functionals. *Trans. Amer. Math. Soc.* **65** 1–13.

[**Kb**] M. Kac (1951). On some connections between probability theory and differential and integral equations, *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability,* University of California Press, Berkeley, 189–215.

[**Ks**] R. Khasminskii (1959). On positive solutions of the equation $Au + Vu = 0$, *Theory of Probability and its Applications,* **4** 309–318.

[**P**] V. G. Papanicolaou (1990). The probabilistic solution of the third boundary value problem for second order elliptic equations. *Probability Theory and Related Fields,* **87** 27–77.

[W] R. J. Williams (1985). A Feynman-Kac gauge for solvability of the Schrödinger equation, *Adv. in Appl. Math.* **6** 1–3.

[Z] Z. Zhao (1983). Conditional gauge with unbounded potential. *Z. Wahrsch. verw. Gebiete,* **65** 13–18.

Ruth J. Williams
Department of Mathematics
University of California, San Diego

This page intentionally left blank

Conditional Brownian motion and conditional gauge by Michael Cranston

Through his works and words, Kai Lai Chung has been the spur for substantial developments in the understanding of conditional Brownian motion and its application to the theory of Schrödinger operators. Many in the field received mail or phone calls from Chung with interesting and provocative questions on the subject. At the Spring 1982 meeting of the Seminar on Stochastic Processes, he posed an interesting question on the lifetime of conditional Brownian motion. The resolution of this question (described below) has led to wide ranging developments. His foundational work with Rao on the gauge theorem, to mention just one of his many works in this area, has served as the motivation for many developments in the understanding of Schrödinger operators and their semigroups. And, of course, "From Brownian Motion to Schrödinger's Equation," with Zhongxin Zhao has served as guide to developments in the field. In this short, semi-accurate, historical note, I'd like to outline a few developments that trace their origins to the encouragement of Chung. I'd like to apologize in advance for the many works which are not mentioned here, due in large part to an interest in brevity.

First, an introduction is in order. Let B denote Brownian motion on $\mathbf{R}^{\mathrm{d}}$ defined on the probability space $(\Omega, \mathcal{F}_t, \{P_x\}_{x\in\mathbf{R}^{\mathrm{d}}})$. For $D \subset \mathbf{R}^{\mathrm{d}}$, let $p(t,x,y)$ be its transition density when killed at time $\tau_D = \inf\{t > 0 : B_t \notin D\}$ (the heat kernel on D.) Given a positive super-harmonic function h on D, define $p^h(t,x,y) = p(t,x,y)h(y)/h(x)$. This is the transition density for a new diffusion called the h-process or conditional Brownian motion. We denote by P_x^h the measure on path space corresponding to Brownian motion started at x and with transition density $p^h(t,x,y)$. In the case when h is the Martin kernel with pole at the Martin boundary point ξ then the conditional Brownian motion exits the domain at the boundary point ξ, in the sense that B_t converges P_x^h *a.s.* to ξ in the Martin topology as t approaches the path lifetime, τ_D. If $h(\cdot) = G_D(\cdot, y)$ where G_D is the Green function for D and $y \in D$, then the $h-$process will converge to y as t approaches the path lifetime. When h is the Martin kernel with pole at ξ we denote the resulting measure by P_x^ξ and by P_x^y when $h(\cdot) = G(\cdot, y)$. These were developments due to Doob [D] in his study of probabilistic versions of the Fatou boundary limit results for harmonic functions. Now it's been known for some time that if D is bounded,

that for unconditioned Brownian motion

$$E_x[\tau_D] < c_d \mathrm{vol}(D)^{2/d}, \tag{1}$$

and if λ_1 is the first Dirichlet eigenvalue for $\frac{1}{2}\Delta$ on D, then

$$\lim_{t\to\infty} \frac{1}{t} \log P_x(\tau_D > t) = -\lambda_1. \tag{2}$$

Using the Martin boundary, denoted here by $\partial_M D$, the expected lifetime can be expressed $E_x[\tau_D] = \int_{\partial_M D} E_x^\xi[\tau_D]\omega_x(d\xi)$, where ω_x is the exit distribution of Brownian motion on $\partial_M D$, also known as harmonic measure. So, by Fubini, $E_x^\xi[\tau_D]$ is finite ω_x almost surely. Chung's question is this:

When is $E_x^\xi[\tau_D]$ bounded uniformly in x and ξ?

Or more generally, when is $E_x^\xi[\tau_D]$ finite? This innocuous sounding question turned out to have quite broad implications. It led to the introduction of some very interesting ideas from analysis into probability theory, such as the boundary Harnack principle, Whitney chains, Littlewood-Paley g-function and intrinsic ultracontractivity.

The first result on this question, due to McConnell and the author [CM], was that there is a positive constant c so that if $D \subset \mathbf{R}^2$, h a positive harmonic function on D, then

$$E_x^h[\tau_D] \le c\mathrm{vol}(D). \tag{3}$$

This is the analog then of (1) in $d = 2$. An example of a bounded $D \subset \mathbf{R}^3$ was given with a $\xi \in \partial_M D$ for which $E_x^\xi[\tau_D] = \infty$. Thus, the analog of (1) can not hold for $d > 2$ without further assumptions. First a word or two on the proof of (3). This relies on decomposing the domain D into subregions by means of the 2^m-level sets of the function h. That is, $D = \bigcup_{m=-\infty}^{\infty} D_m$, where $D_m = \{x \in D : 2^{m-1} < h(x) < 2^{m+1}\}$. The conditional Brownian motion viewed at the successive hitting times to $C_m = \{x \in D : h(x) = 2^m\}$ forms a birth and death Markov chain on $\{2^m : m \in \mathbf{Z}\}$ with probability 2/3 of going up and 1/3 of going down. This implies the number of visits to the C_m are geometrically distributed random variables. These have finite expectation with a value independent of m. The other key observation is that the expected amount of time the conditional Brownian motion spends in D_m starting from C_m is equivalent (owing to the fact that $1/2 \le h(y)/h(x) \le 2, x \in C_m, y \in D_m$) to the amount of time standard Brownian motion spends in D_m starting from C_m. Combining this observation with (1) gives that the expected time spent in D_m starting on C_m by the h-process is bounded by $C_d \mathrm{vol}(D_m)^{2/d}$. Using the strong Markov property and summing leads to an upper bound of

$$E_x^h[\tau_D] \le C_d \sum_{m=-\infty}^{\infty} \mathrm{vol}(D_m)^{2/d}.$$

In case $d = 2$, the sum is bounded by $2\text{vol}(D)$ leading to the result that there is a constant c such that (3) holds for $D \subset \mathbf{R}^2, x \in D, h$ any positive superharmonic on D. Since $2/d < 1$ for $d \geq 3$, the finiteness of $\sum_{m=-\infty}^{\infty} \text{vol}(D_m)^{2/d}$ does not generally hold and leads to interesting questions about the influence of the regularity of the boundary and its effect on the size of the sets D_m. (The relation between boundary regularity and the growth of harmonic functions is a key issue in the subject.) This question was addressed by Bañuelos [B], Falkner [F1], Bass and Burdzy [BB1], De Blassie [DeB], Kenig and Pipher [KP], and myself [C1], among others in the higher dimensional case. The results of Bañuelos [B] incorporated many of the types of domains encountered in analysis, namely Lipschitz, NTA (non-tangentially accessible), John and BMO-extension (uniform) domains. In order to describe the results in [B], we consider a Whitney decomposition of D. This is a collection of closed squares Q_j with sides parallel to the coordinate axes and $D = \bigcup_j Q_j$ with the properties:

$$Q_j^o \bigcap Q_k^o = \emptyset, j \neq k,$$
$$\frac{1}{4} \leq \frac{l(Q_j)}{l(Q_k)} \leq 4, \text{ if} Q_j \bigcap Q_k \neq \emptyset,$$
$$1 \leq \frac{d(Q_j, \partial D)}{l(Q_j)} \leq 4\sqrt{d}, \text{ for all} j.$$

A Whitney chain connecting Q_j and Q_k is a sequence of Whitney squares $\{Q_{m_i}\}_{i=0}^n$ with $Q_{m_0} = Q_j$, $Q_{m_n} = Q_k$ and $Q_{m_i} \bigcup Q_{m_{i+1}} \neq \emptyset$. An important fact about Whitney squares is that there is a positive constant c so that for any positive harmonic function h in D and adjacent Whitney squares $Q_j \bigcup Q_k \neq \emptyset$, we have $h(x) < ch(y), x \in Q_j, y \in Q_k$. Whitney chains are very well suited to the study of conditional Brownian motion. The reason is that due to Harnack's inequality, any positive harmonic function will be 'flat' on the Whitney square Q_j. This means that the transition densities $p^h(t,x,y)$ and $p(t,x,y)$ will be equivalent on Q_j which means the behavior of ordinary and conditional Brownian motion will be comparable on Q_j. Now define the quasi-hyperbolic distance from $x \in Q_j$ to x_0 by first setting $d(x) = dist(x, \partial D)$ and then putting,

$$\rho_D(x_1, x_2) = \inf_\gamma \int_\gamma \frac{ds}{d(\gamma(s))},$$

with the inf being taken over all rectifiable curves in D from x_1 to x_2. Taking points $x_1 \in Q_j, x_2 \in Q_k$, we have

$$\rho_D(x_1, x_2) \approx \text{length of shortest Whitney chain from } Q_j \text{ to } Q_k.$$

Note that repeated applications of Harnack's inequality in successive squares in a Whitney chain implies that $h(x_1) \leq c^{\rho_D(x_1,x_2)} h(x_2)$. If we fix an $x_0 \in D$ and write $\rho_D(x) = \rho_D(x, x_0)$, this implies that for some constant C,

$$D_m \subset \{x \in D : \rho_D(x) > C|m|\}.$$

Now a result of Smith and Stegenga [SS] implies that for a class of domains called Hölder of order 0 , $H(0)$ (this class includes Lipschitz, NTA, John and BMO-extension domains) one has $\rho_D(\cdot, x_0) \in L^p(D)$ for any $0 < p < \infty$. Using this, Bañuelos obtains for D a bounded Hölder of order 0 domain that

$$\sum_{m=-\infty}^{\infty} \mathrm{vol}(D_m)^{2/d} < \infty.$$

This implies that $H(0)$ domains are regular enough so that an analog of (3) holds for them in all dimensions. There are also beautiful connections in simply connected planar domains between the behavior of conditional Brownian motion and the hyperbolic geometry of the region. This was developed in Bañuelos and Carrol [BC] and Davis [D]. We start our exposition of this connection with an observation of Bañuelos [B1]. If D is a simply connected, planar domain, and $\phi : B(0,1) \to D$ maps the unit disc $B(0,1)$ of the complex plane conformally onto D with $\phi(0) = x$, then

$$g_*^2(\phi)(\theta) = \frac{1}{\pi} \int_{B(0,1)} \log\left(\frac{1}{|z|}\right) \frac{1-|z|^2}{|z-e^{i\theta}|^2} |\phi'(z)|^2 dz$$

is the Littlewood-Paley square function. Recalling that the Green function of $B(0,1)$ with pole at the origin is $\log(\frac{1}{|z|})$ and that the Green function is preserved by conformal mappings, it is easy to deduce that for h a positive harmonic function on D with the representation

$$h(z) = \int_0^{2\pi} \frac{1-|z|^2}{|z-e^{i\theta}|} d\mu(\theta)$$

for a positive Borel measure μ on $\partial B(0,1)$, that

$$\begin{aligned} E_x^h[\tau_D] &= \frac{1}{h(x)} \int_D G_D(x,y) h(y) dy \qquad (4) \\ &= \frac{1}{h(x)} \int_{B(0,1)} \log(\frac{1}{|z|}) h(\phi(z)) |\phi'(z)|^2 dx dy \\ &= \frac{1}{h(x)} \int_0^{2\pi} g_*^2(\phi)(\theta) d\mu(\theta). \end{aligned}$$

Since $\mu([0,2\pi]) = h(x)$ and

$$g_*^2(\phi)(\theta) \leq C \int_{B(0,1)} |\phi'(z)|^2 dz \leq C\mathrm{vol}(D)$$

it follows that

$$E_x^h[\tau_D] \leq C\mathrm{vol}(D)$$

giving another derivation of the lifetime estimate in the special case of simply connected planar domains. But this gives additional information as developed in Bañuelos and Carrol [BC]. There the authors observed that if $K(z,\xi)$ is the Poisson

kernel for $B(0,1)$ with pole at $\xi \in \partial B(0,1)$ then there are positive constants c, C such that

$$c \sup_{\phi} g_*^2(\phi)(0) \leq \sup_{\phi} \int_{B(0,1)} K(z,1)K(z,-1)|\phi'(z)|^2 dxdy \leq C \sup_{\phi} g_*^2(\phi)(0),$$

where the sup is taken over all conformal mappings $\phi : B(0,1) \to D$ with $\phi(0) = x$. But another equivalence holds for $K(z,1)K(z,-1)$. Namely, if $d(z,\Gamma)$ denotes the hyperbolic distance in $B(0,1)$ from z to the geodesic $\tilde{\Gamma} = [-1,1]$ then

$$\frac{1}{4} K(z,1)K(z,-1) \leq e^{-2d(z,\tilde{\Gamma})} \leq K(z,1)K(z,-1).$$

Using the conformal invariance of the hyperbolic metric, writing d_D for the hyperbolic metric in D and putting these two equivalences together yields the existence of two positive constants c, C such that

$$c \sup_{\Gamma} \int_D e^{-2d_D(z,\Gamma)} \leq sup_{x,h} E_x^h[\tau_D] \leq C \sup_{\Gamma} \int_D e^{-2d_D(z,\Gamma)}.$$

This has a beautiful Corollary involving the Whitney decomposition mentioned above. Let Q be a Whitney cube with center z_Q and let T_Q be the total amount of time spent in Q before τ_D. Then for Martin boundary points ξ_1, ξ_2 and Γ the hyperbolic geodesic connecting them, there are positive constants c, C such that

$$\frac{1}{4} e^{-Cd_D(z_Q,\Gamma)} \leq E_{\xi_1}^{\xi_2}[T_Q] \leq e^{-cd_D(z_Q,\Gamma)}.$$

This is a quantitative statement about how closely the conditional Brownian motion from ξ_1 to ξ_2 follows the hyperbolic geodesic from ξ_1 to ξ_2. Davis [D1] pursued this connection further in estimating the variance of τ_D under the measure $E_{\xi_1}^{\xi_2}$. If Q and R are Whitney squares and $P_Q = P_{\xi_1}^{\xi_2}(\tau_{D \cap Q^c} < \tau_D)$ with a similar definition for P_R and letting $\delta(D)$ be the area of the largest disc which can be inscribed in D, then

$$|Cov_{\xi_1}^{\xi_2}(T_Q, T_R)| \leq Ce^{-c\delta_D(z_Q,z_R)} \mathrm{vol}(Q)\mathrm{vol}(R)(P_Q + P_R),$$

and

$$Var_{\xi_1}^{\xi_2}(\tau_D) \leq \delta(D) E_{\xi_1}^{\xi_2}[\tau_D].$$

The first of these shows exactly how the decay of the dependence between the occupation times T_Q and T_R depends on the hyperbolic distance between Q and R. The second confirms the intuition that the conditional Brownian motion speeds up when traversing narrow channels (take D to be a rectangle of length n and width $\frac{1}{n}$, then for ξ_1 and ξ_2 on opposite ends of the long side of the rectangle, $E_{\xi_1}^{\xi_2}[\tau_D] \leq c$ and $Var_{\xi_1}^{\xi_2}[\tau_D] \leq \frac{c}{n}$. Thus, the conditional motion must go a distance n in a time with bounded expectation, independent of n, but with variance bounded by $\frac{1}{n}$. This means the mass of the measure $P_{\xi_1}^{\xi_2}$ is concentrating on paths which make the length n trip in a time which is some constant that doesn't depend on n.) Refinements and further progress in these directions can be found in the works of

Griffin, McConnell, Verchota [GMV], Griffin, Verchota, Vogel [GVV], Zhang [Zh], Davis and Zhang [DZ], and Xu [X], to name but a few.

Now let's turn our attention to the problem of deciding to what extent the analog of (2) holds for conditional Brownian motion. From the case of a ball $D = \{x : |x| < r\}$ in Euclidean space where $P_0(\tau_D > t) = P_0^\xi(\tau_D > t)$ for every boundary point ξ one might suspect that with some smoothness in $d > 2$ and maybe even with $\mathrm{vol}(D) < \infty$ in $d = 2$ that if $H^+(D)$ is the class of positive harmonic functions on D, then

$$\lim_{t\to\infty} \frac{1}{t} \log P_x^h(\tau_D > t) = -\lambda_1,\ x \in D\,, h \in H^+(D.) \tag{5}$$

This was addressed in De Blassie [DeB] where it was proved that (6) holds provided D is a Lipschitz domain with sufficiently small Lipschitz constant. Later, Kenig and Pipher [KP], extended this result to Lipschitz domains and NTA domains. Perhaps the nicest approach is due to Bañuelos [B2] and Bañuelos and Davis [BD], which illuminates the relation between the tail behavior of the lifetime of conditional Brownian motion and intrinsic ultracontractivity. The notion of intrinsic ultracontractivity is stipulated in Davies and Simon [DS], as the property that the semigroup of the ground state transformation of an operator maps L^2 to L^∞. To make this definition precise in the current setting, if φ_1 is the first Dirichlet eigenfunction for $\frac{1}{2}\Delta$ on D, define a semigroup on $L^2(\varphi_1^2 dx)$ by

$$P_t^{\varphi_1} f(x) = \int_D \frac{e^{\lambda_1 t} p(t,x,y)}{\varphi_1(x)\phi_1(y)} f(y)\varphi_1^2(y)dy,\ f \in L^2(\varphi_1^2 dx).$$

Then the domain D is defined to be intrinsically ultracontractive (IU) if there exist constants C_t such that

$$|P_t^{\varphi_1} f(x)| \le C_t ||f||_{L^2(\varphi_1^2 dx)},\ t > 0$$

An important consequence of IU is that for any $\epsilon > 0$ there is a $t(\epsilon)$ such that

$$(1-\epsilon)e^{-\lambda_1 t}\varphi_1(x)\varphi_1(y) \le p(t,x,y) \le (1+\epsilon)e^{-\lambda_1 t}\varphi_1(x)\varphi_1(y). \tag{6}$$

Since, for any $h \in H^+(D)$,

$$P_x^h(\tau_D \ge t) = \frac{1}{h(x)} \int_D p(t,x,y)h(y)dy \le 1,$$

it follows easily from (5) that

$$\lim_{t\to\infty} \frac{1}{t} \log P_x^h(\tau_D \ge t) = -\lambda_1,$$

giving the Bañuelos analog of (2) for conditional Brownian motion on IU domains. In the case of planar domains of finite area, Bañuelos and Davis [BD] proved the following analog of IU :for each $x \in D$,

$$\lim_{t\to\infty} \frac{e^{\lambda_1 t} p(t,x,y)}{\varphi_1(x)\varphi_1(y)} = 1, \text{ uniformly in } y \in D.$$

This implies that the analog of (2) for conditional Brownian motion holds for planar domains of finite area.

Another application of conditional Brownian motion, which has been an area of research to which Professor Chung has made many contributions, is to the study of the Schrödinger equation by means of the Feynman-Kac formula. A seminal paper on the subject was that of Aizenman and Simon [AS], who used path integral techniques (the Feynman-Kac formula) to prove Harnack's inequality for Schrödinger operators. Consider, with $d > 2$ for ease of presentation, a potential V satisfying

$$\lim_{r\to 0} \sup_{x\in \mathbf{R}^d} \int_{|x-y|<r} \frac{|V(y)|}{|x-y|^{d-2}} dy = 0.$$

The class of such potentials is called the Kato class and is denoted by K_d. They are particularly well suited to the Newtonian potential and thus as well to the occupation properties of Brownian motion. Now, consider the Dirichlet problem, for $f \in C(\partial D)$,

$$\begin{aligned} \frac{1}{2}\Delta u(x) + V(x)u(x) = 0,\ x \in D, \\ u(x) = f(x),\ x \in \partial D. \end{aligned} \tag{7}$$

The Gauge Theorem of Chung and Rao (see the article of Ruth Williams in this volume) says that either $E_x[e^{\int_0^{\tau_D} V(B_s)ds}] \equiv \infty$ on D or this quantity, called the gauge, is bounded on D. Let's assume that the second alternative of this dichotomy holds. Then by Feynman-Kac, the solution of (6) is then given by

$$u(x) = E_x[e^{\int_0^{\tau_D} V(B_s)ds} f(B_{\tau_D})].$$

Let's suppose now that D is a Lipschitz domain so that the Euclidean and Martin boundary of D are the same. Then decompose the Feynman-Kac formula using conditional Brownian motion,

$$u(x) = \int_{\partial D} E_x^y[e^{\int_0^{\tau_D} V(B_s)ds}] f(y) P_x(B_{\tau_D} \in dy). \tag{8}$$

Now the analog of Chung's question regarding the finiteness of the expected lifetime of conditional Brownian motion as well as his question regarding the finiteness of the gauge is when is

$$E_x^y[e^{\int_0^{\tau_D} V(B_s)ds}] < \infty?$$

The quantity $u(x,y) = E_x^y[e^{\int_0^{\tau_D} V(B_s)ds}]$ is known as the conditional gauge. Under the conditions set down above, namely that D be a Lipschitz domain and $V \in K_d$ a dichotomy, similar to the Gauge Theorem, holds: either

$$E_x^y[e^{\int_0^{\tau_D} V(B_s)ds}] \equiv \infty$$

or there are positive constants c, C such that

$$c \le E_x^y[e^{\int_0^{\tau_D} V(B_s)ds}] \le C, \text{ for all } x, y \in D \bigcup \partial D. \tag{9}$$

This is called the Conditional Gauge Theorem (CGT). It can be viewed as a statement on the mixing properties of conditional Brownian motion. The potential V may possess singularities. The CGT says that these singularities can't be so bad that P_x^y-paths would miss them in the sense that $E_x^y[e^{\int_0^{\tau_D} V(B_s)ds}] < \infty$ for one pair x, y but for another pair of points, z, w, one has $E_x^y[e^{\int_0^{\tau_D} V(B_s)ds}] = \infty$. That is under both measures, P_x^y and P_z^w, the occupation distributions of paths are similar enough that they will simultaneously give a finite answer or an infinite answer when asked about the value of the conditional gauge. This requires some smoothness of ∂D with its resulting effect on the behavior of the Green function. Early results on the subject were those of Falkner [F2] and Zhao [Z1], [Z2]. In the fundamental works of Zhao, the CGT was proved for Kato class potentials on the ball and then domains with C^2 boundary. For Lipschitz domains and Kato potentials, the result was proven in Cranston, Fabes and Zhao [CFZ]. The extension to Lipschitz domains of the CGT used the so-called $3G$-Theorem. This result says that if G is the Green function for $\frac{1}{2}\Delta$ on D, then there is a positive constant C such that

$$\frac{G(x,z)G(z,y)}{G(x,y)} \leq C\left(\frac{1}{|x-z|^{d-2}} + \frac{1}{|y-z|^{d-2}}\right). \tag{10}$$

The left hand side in (10) is the Green function for conditional Brownian motion started at x and conditioned to exit D at y. This is the occupation density for conditional Brownian motion in D, in the sense that the total expected amount of time spent by B in $A \subset D$ with respect to the measure P_x^y is $\int_A \frac{G(x,z)G(z,y)}{G(x,y)}dz$. The right hand side of (10) is the sum of the Newtonian potentials with poles at x and y , respectively. These are the occupation densities for unconditioned Brownian motion in $\mathbf{R^d}$ started at x and y. The $3G$-inequality says that if $V \in K_d$ and is thus well adapted to the occupation measure of Brownian motion (unconditioned) then it is also well adapted to the occupation measure of conditional Brownian motion. In the case when the conditional gauge is finite, the CGT permits comparisons between potential theoretic quantities for the two operators $-\frac{1}{2}\Delta$ and $-\frac{1}{2}\Delta + V$. This lies close to the original motivation of Aizenman and Simon [AS]. For example, suppose that for some $f \in C(\partial D)$,

$$\frac{1}{2}\Delta v(x) = 0, x \in D,\ v(x) = f(x),\ x \in \partial D.$$

and

$$-\frac{1}{2}\Delta u(x) + V(x)u(x) = 0, x \in D,\ u(x) = f(x),\ x \in \partial D.$$

Then $v(x) = E_x[f(B_{\tau_D})]$ and since $c \leq E_x^y[e^{\int_0^{\tau_D} V(B_s)ds}] \leq C$ it follows from (4) that $cv(x) \leq u(x) \leq Cv(x)$, $x \in D$. With this equivalence, Harnack's inequality and even the boundary Harnack inequality can be deduced for positive solutions in D of $-\frac{1}{2}\Delta u(x) + V(x)ux = 0$. Many other similar conclusions follow in an equally easy manner. Using the simple formula

$$u(x,y) = \frac{G_V(x,y)}{G(x,y)}, \tag{11}$$

it follows that

$$cG(x,y) \leq G_V(x,y) \leq CG(x,y),\ x,y \in D,$$

where G_V is the Green function for $-\frac{1}{2}\Delta + V$. Since the Martin kernels $K(x,\xi)$ and $K_V(x,\xi)$ are the limit of ratios of the Green functions $G(x,\xi)$ and $G_V(x,\xi)$ it follows as well that

$$cK(x,y) \leq K_V(x,y) \leq CK(x,y),\ x,y \in D,$$

where K and K_V are the Martin kernels for $\frac{1}{2}\Delta$ and $-\frac{1}{2}\Delta + V$, respectively. Two dimensional versions of these results appeared in Bass and Burdzy [BB2], Cranston [C2], McConnell [M], and Zhao [Z3]. Results similar in flavor and which also incorporate the notion of IU above are due to Bañuelos [B1] who proved that when the conditional gauge is finite and D is a Lipschitz or NTA domain that there exist positive constants c_t, C_t such that

$$c_t p(t,x,y) \leq p_V(t,x,y) \leq C_t p(t,x,y).,\ t > 0,\ x,y \in D,$$

where p_V is the heat kernel for $-\frac{1}{2}\Delta + V$. An additional result of Bañuelos in this connection is that if the conditional gauge is finite and D is an $H(0)$ domain, as described earlier, then the operator $-\frac{1}{2}\Delta + V$ is IU. It's interesting to note that the proofs used log Sobolev inequalities. Further developments appear in a series of papers by Chen and Song [ChS1], [ChS2], and Chen [Ch] among others. In [ChS1], the authors follow the developments of Bañuelos [B1], and consider the conditional gauge problem for the fractional Laplacian, $(-\Delta)^\alpha$ for $0 < \alpha < 2$ and potentials in the suitably modified Kato class $K_{\alpha,d}$ where $V \in K_{\alpha,d}$ if

$$\lim_{r\to 0} \sup_{\{x \in \mathbf{R^d}\}} \int_{|x-y|<r} \frac{|V(y)|}{|x-y|^{d-\alpha}}\, dy = 0.$$

In this paper, Chen and Song deduced the CGT on $C^{1,1}$ domains for the operator $(-\Delta)^\alpha$ and $K_{\alpha,d}$ potentials. The proper process to use in the Feynman-Kac representation in this case is the symmetric stable process of order α ,X, rather than the Brownian motion used when considering Δ. Their approach was to split the potential writing $V = V_1 + V_2$, where $V_2 \in \mathbf{L}^\infty$ and V_1 has a small Kato norm, $\sup_{x\in D} \int_D \frac{|V_1(y)|}{|x-y|^{d-\alpha}} dy$ is small. Then by a simple lemma of Khasminski, they show that the Green functions, $G^\alpha_{V_1}$, for $(-\Delta)^\alpha + V_1$ and G^α for $(-\Delta)^\alpha$ on D satisfy $G^\alpha_{V_1} \approx G^\alpha$ in the sense that there are positive constants c, C such that $cG^\alpha \leq G^\alpha_{V_1} \leq CG^\alpha$. This equivalence can then be used to prove that $(-\Delta)^\alpha + V$ is IU. However, $(-\Delta)^\alpha + V$ being IU implies that $G^\alpha \approx G^\alpha_V$. Now using the formula the analog of (7), the finiteness of the right hand side follows from the inequality $G^\alpha_{V_1} \leq CG^\alpha$ and the $3G$-Theorem for the Green function G^α on D, the CGT follows. This was extended in [ChS2] to $H(0)$ domains again by an approach inspired by [B1].

Relations between subcriticality and boundedness of the conditional gauge have been investigated by Zhao [Z4]. Consider the class $B_c = \{q : \mathbf{R^d} \to \mathbf{R} :$ *supp* q compact$\} \bigcap \mathbf{L}^\infty$. The operator $-\frac{1}{2}\Delta + V$ is called subcritical if

$$\forall q \in B_c,\ \exists \epsilon > 0, \text{ such that } -\frac{1}{2}\Delta + V + \epsilon q \geq 0.$$

This amounts to a strict positivity of $-\frac{1}{2}\Delta + V$.

For a subclass of K_d potentials which satisfy a condition at ∞, we refer the reader to [Z4] for the details, Zhao proved that subcriticality is equivalent to

$$u(x,y) = E_x^y[e^{\int_0^\infty V(B_s)ds}] \text{ is bounded on } \mathbf{R^d} \times \mathbf{R^d}.$$

There were many other equivalences in that work which go a long way toward establishing the power of the approach in investigating the Schrödinger operator.

Generalizations of the conditional gauge theorem to broader classes of Markov processes and potentials including measures, have been carried out in Chen and Song [ChS5] and Chen [Ch]. In the last work, Chen has proved Gauge and Conditional Gauge Theorems for a new class of Kato potentials, which even includes singular measures, and general transient Borel right processes. And most strikingly, following a suggestion of Chung, that the CGT is actually the Gauge Theorem for the conditional process!

In this review, we've examined some of the many results which have connections with the works of Chung to be found in this volume. While we haven't explicitly drawn the connections, we hope that these ties will become obvious to any reader of this volume. Finally, the author would like to express his gratitude to Professor Chung for introducing him to the fascinating problems in this area.

References

[AS] Aizenman, M.; Simon, B., Brownian motion and Harnack inequality for Schrödinger operators. *Comm. Pure Appl. Math.* 35 (1982), no. 2, 209–273.

[B] Bañuelos, Rodrigo, On an estimate of Cranston and McConnell for elliptic diffusions in uniform domains. *Probab. Theory Related Fields* 76 (1987), no. 3, 311–323.

[B1] Bañuelos, Rodrigo, Lifetime and heat kernel estimates in nonsmooth domains. *Partial differential equations with minimal smoothness and applications* (Chicago, IL, 1990), 37–48, IMA Vol. Math. Appl., 42, Springer, New York, 1992.

[B2] Bañuelos, Rodrigo, Intrinsic ultracontractivity and eigenfunction estimates for Schrödinger operators. *J. Funct. Anal.* 100 (1991), no. 1, 181–206.

[BC] Bañuelos, Rodrigo; Carroll, Tom, Conditioned Brownian motion and hyperbolic geodesics in simply connected domains. *Michigan Math. J.* 40 (1993), no. 2, 321–332.

[BD] Bañuelos, Rodrigo; Davis, Burgess, Heat kernel, eigenfunctions, and conditioned Brownian motion in planar domains. *J. Funct. Anal.* 84 (1989), no. 1, 188–200.

[BB1] Bass, Richard F.; Burdzy, Krzysztof, Lifetimes of conditioned diffusions. *Probab. Theory Related Fields* 91 (1992), no. 3-4, 405–443.

[BB2] Bass, Richard F.; Burdzy, Krzysztof, Conditioned Brownian motion in planar domains. *Probab. Theory Related Fields* 101 (1995), no. 4, 479–493.

[Ch] Chen, Zhen-Qing, Gaugeability and conditional gaugeability. *Trans. Amer. Math. Soc.* 354 (2002), no. 11, 4639–4679.

[ChS1] Chen, Zhen-Qing; Song, Renming, General gauge and conditional gauge theorems. *Ann. Probab.* 30 (2002), no. 3, 1313–1339.

[ChS2] Chen, Zhen-Qing; Song, Renming, Conditional gauge theorem for non-local Feynman-Kac transforms. *Probab. Theory Related Fields* 125 (2003), no. 1, 45–72.

[ChS3] Chen, Zhen-Qing; Song, Renming, Intrinsic ultracontractivity, conditional lifetimes and conditional gauge for symmetric stable processes on rough domains. *Illinois J. Math.* 44 (2000), no. 1, 138–160.

[ChS4] Chen, Zhen-Qing; Song, Renming, Intrinsic ultracontractivity and conditional gauge for symmetric stable processes. *J. Funct. Anal.* 150 (1997), no. 1, 204–239.

[ChS5] Chen, Zhen-Qing; Song, Renming, General gauge and conditional gauge theorems. *Ann. Prob.* 30 (2002), no,3, 1313–1339.

[ChZ] Chen, Zhen-Qing; Zhang, Tu-Sheng, Girsanov and Feynman-Kac type transformations for symmetric Markov processes. *Ann. Inst. H. Poincar Probab. Statist.* 38 (2002),

[C2] Cranston, M., Conditional Brownian motion, Whitney squares and the conditional gauge theorem. *Seminar on Stochastic Processes,* 1988 (Gainesville, FL, 1988), 109–119, Progr. Probab., 17, Birkhuser Boston, Boston, MA, 1989.

[C1] Cranston, M., Lifetime of conditioned Brownian motion in Lipschitz domains. *Z. Wahrsch. Verw. Gebiete* 70 (1985), no. 3, 335–340.

[CFZ] Cranston, M.; Fabes, E.; Zhao, Z., Conditional gauge and potential theory for the Schrödinger operator. *Trans. Amer. Math. Soc.* 307 (1988), no. 1, 171–194.

[CM] Cranston, M.; McConnell, T., The lifetime of conditioned Brownian motion in the plane. *Zeit. Wahr. Verw. Geb.* 65, (1983),1-11.

[D1] Davis, Burgess, Conditioned Brownian motion in planar domains. *Duke Math. J.* 57 (1988), no. 2, 397–421.

[DZ] Davis, B,; Zhang, B., Moments of the lifetime of conditioned Brownian motion in cones. *Proc. Amer. Math. Soc.* 121 (1994), no. 3, 925–929.

[DS] Davies, E. B.; Simon, B., L^1-properties of intrinsic Schrödinger semigroups. *J. Funct. Anal.* 65 (1986), no. 1, 126–146.

[DeB] De Blassie, R. Dante, Doob's conditioned diffusions and their lifetimes. *Ann. Probab.*

[D] Doob, J. L., Conditional Brownian motion and the boundary limits of harmonic functions. *Bull. Soc. Math. France*, 85:431–458, 1957.

[F1] Falkner, Neil, Conditional Brownian motion in rapidly exhaustible domains.

Ann. Probab. 15 (1987), no. 4, 1501–1514.

[F2] Falkner, Neil, Feynman-Kac functionals and positive solutions of $\frac{1}{2}\Delta u + qu = 0$. *Z. Wahrsch. Verw. Gebiete* 65 (1983), no. 1, 19–33.

[GMV] Griffin, Philip S.; McConnell, Terry R.; Verchota, Gregory, Conditioned Brownian motion in simply connected planar domains. *Ann. Inst. H. Poincaré Probab. Statist.* 29 (1993), no. 2, 229–249.

[GVV] Griffin, P. S.; Verchota, G. C.; Vogel, A. L., Distortion of area and conditioned Brownian motion. *Probab. Theory Related Fields* 96 (1993), no. 3, 385–413.

[KP] Kenig, Carlos E.; Pipher, Jill, The h-path distribution of the lifetime of conditioned Brownian motion for nonsmooth domains. *Probab. Theory Related Fields* 82 (1989), no. 4, 615–623.

[M] McConnell, Terry R., A conformal inequality related to the conditional gauge theorem. *Amer. Math. Soc.Trans.*

[SS] Smith, W.; Stegenga, D., Hölder domains and Poincaré domains. *Trans. Amer. Math. Soc.* 319 (1990), no. 1, 67–100.

[X] Xu, J., The lifetime of conditioned Brownian motion in domains of infinite area. *Probab. Theory Related Fields* 87 (1990), 469-487.

[Zh] Zhang, Biao, On the variances of occupation times of conditioned Brownian motion. *Trans. Amer. Math. Soc.* 348 (1996), no. 1, 173–185.

[Z1] Zhao, Zhongxin, Conditional gauge with unbounded potential. *Z. Wahrsch. Verw. Gebiete* 65 (1983), no. 1, 13–18.

[Z2] Zhao, Zhongxin, Uniform boundedness of conditional gauge and Schrödinger equations. *Comm. Math. Phys.* 93 (1984), no. 1, 19–31.

[Z3] Zhao, Zhongxin, Green functions and conditioned gauge theorem for a 2-dimensional domain. *Seminar on Stochastic Processes*, 1987 (Princeton, NJ, 1987), 283–294, Progr. Probab. Statist., 15, Birkhuser Boston, Boston, MA, 1988.

[Z4] Zhao, Zhongxin, Subcriticality and gaugeability of the Schrödinger operator. *Trans. Amer. Math. Soc.* 334 (1992), no. 1, 75–96.

[Z5] Zhao, Zhongxin, A probabilistic principle and generalized Schrödinger perturbation. *J. Funct. Anal.* 101 (1991), no. 1, 162–176.

Michael Cranston
Department of Mathematics
University of California, Irvine

PART 3
Selected Works of Kai Lai Chung

Reprinted from *Ann. Math. Stat.* 12 (1941), pp. 328–338
with permission from Institute of Mathematical Statistics.

ON THE PROBABILITY OF THE OCCURRENCE OF AT LEAST m EVENTS AMONG n ARBITRARY EVENTS

By Kai Lai Chung

Tsing Hua University, Kunming, China

Introduction. Let $E_1, \cdots, E_n$, denote n arbitrary events. Let $p_{\nu_1' \cdots \nu_i' \nu_{i+1} \cdots \nu_j}$, where $0 \leqq i \leqq j \leqq n$ and $(\nu_1, \cdots, \nu_j)$ is a combination of the integers $(1, \cdots, n)$, denote the probability of the non-occurrence of $E_{\nu_1}, \cdots, E_{\nu_i}$ and the occurrence of $E_{\nu_{i+1}}, \cdots, E_{\nu_j}$. Let $p_{[\nu_1 \cdots \nu_i]}$ denote the probability of the occurrence of $E_{\nu_1}, \cdots, E_{\nu_i}$ and no others among the n events. Let $S_j = \Sigma p_{\nu_1 \cdots \nu_j}$ where the summation extends to all combinations of j of the n integers $(1, \cdots, n)$. Let $p_m(\nu_1, \cdots, \nu_k)$, $(1 \leqq m \leqq k \leqq n)$, denote the probability of the occurrence of at least m events among the k events $E_{\nu_1}, \cdots, E_{\nu_k}$.

By the set $(x_1, \cdots, x_b, \cdots, x_a) - (x_1, \cdots, x_b)$ (where $b \leqq a$) we mean the set $(x_{b+1}, \cdots, x_a)$. And by a $\binom{a}{b}$-combination out of $(x_1, \cdots, x_a)$ we mean a combination of b integers out of the a integers $(x_1, \cdots, x_a)$.

We often use summation signs with their meaning understood, thus for a fixed k, $1 \leqq k \leqq n$, the summations in $\Sigma p_{\nu_1 \cdots \nu_k}$, or $\Sigma p_m(\nu_1, \cdots, \nu_k)$, extend to all the $\binom{n}{k}$-combinations out of $(1, \cdots, n)$.

The following conventions concerning the binomial coefficients are made:

$$\binom{0}{0} = 1, \qquad \binom{a}{b} = 0 \qquad \text{if} \qquad a < b \qquad \text{or if} \qquad b < 0.$$

It is a fundamental theorem in the theory of probability that, if $E_1, \cdots, E_n$ are incompatible (or "mutually exclusive"), then

$$p_1(1, \cdots, n) = p_1 + \cdots + p_n .$$

When the events are arbitrary, we have Boole's inequality

$$p_1(1, \cdots, n) \leqq p_1 + \cdots + p_n .$$

Gumbel[1] has generalized this inequality to the following:

$$p_1(1, \cdots, n) \leqq \frac{\Sigma p_1(\nu_1, \cdots, \nu_k)}{\binom{n-1}{k-1}},$$

[1] *C. R. Acad. Sc.* Vol. 205(1937), p. 774.

for $k = 1, \cdots, n$. The case $k = 1$ gives Boole's inequality. Fréchet[2] has announced that Gumbel's result can be sharpened to the following

$$A_{k+1} = \frac{\Sigma p_1(\nu_1, \cdots, \nu_{k+1})}{\binom{n-1}{k}} \leqq \frac{\Sigma p_1(\nu_1, \cdots, \nu_k)}{\binom{n-1}{k-1}} = A_k, \tag{1}$$

for $k = 1, \cdots, n - 1$. Thus, A_k is non-increasing for k increasing. On the other hand, Poincaré has obtained the following formula which expresses $p_1(1, \cdots, n)$ in terms of the S_j's,

$$\begin{aligned} p_1(1, \cdots, n) = \sum p_{\nu_1} - \sum p_{\nu_1\nu_2} + \sum p_{\nu_1\nu_2\nu_3} - \cdots \\ + (-1)^n p_{1\cdots n} = \sum_{j=1}^{n} (-1)^{j-1} S_j. \end{aligned} \tag{2}$$

In the present paper we shall study the more general function $p_m(\nu_1, \cdots, \nu_k)$ as defined above. First we generalize Poincaré's formula and Fréchet's inequalities. In Theorem 1 we establish (for $1 \leqq m \leqq n$)

$$\begin{aligned} p_m(1, \cdots, n) &= \sum p_{\nu_1\cdots\nu_m} - \binom{m}{1} \sum p_{\nu_1\cdots\nu_{m+1}} \\ &\quad + \binom{m+1}{2} \sum p_{\nu_1\cdots\nu_{m+2}} + \cdots + (-1)^{n-m} \binom{n-1}{m-1} p_{1\cdots n} \\ &= \sum_{i=0}^{n-m} (-1)^i \binom{m+i-1}{i} S_{m+i}. \end{aligned} \tag{3}$$

Although this result is well known, we prove it in preparation for Theorem 2. Theorem 3 establishes

$$A_{k+1}^{(m)} = \frac{\Sigma p_m(\nu_1, \cdots, \nu_{k+1})}{\binom{n-m}{k+1-m}} \leqq \frac{\Sigma p_m(\nu_1, \cdots, \nu_k)}{\binom{n-m}{k-m}} = A_k^{(m)}, \tag{4}$$

for $k = 1, \cdots, n - 1$ and $1 \leqq m \leqq k$.

Next, we extend the inequalities (4), and in Theorem 4 we show that

$$A_k^{(m)} \leqq \tfrac{1}{2}(A_{k-1}^{(m)} + A_{k+1}^{(m)}); \tag{5}$$

which states that the differences $A_k - A_{k+1}$ $(k = 1, \cdots, n - 1)$ are non-decreasing for increasing k. From this and a simple result we can deduce (4). Also Theorem 2 establishes that

$$\sum_{i=0}^{2l+1} (-1)^i \binom{m+i-1}{i} S_{m+i} \leqq p_m(1, \cdots, n) \leqq \sum_{i=0}^{2l} (-1)^i \binom{m+i-1}{i} S_{m+i}, \tag{6}$$

[2] Loc. cit., Vol. 208(1939), p. 1703.

for $2l + 1 \leqq n - m$ and $2l \leqq n - m$ respectively. These inequalities throw light on formula (3) and are sharper than the following analogue of Boole's inequality for $p_m(1, \cdots, n)$, which is a special case of (4):

$$p_m(1, \cdots, n) \leqq \Sigma p_{\nu_1 \cdots \nu_m}. \tag{7}$$

The last statement will be evident in the proof.

In Theorem 5 we give an "inversion" of the formula (3), i.e. we express $p_{1 \cdots n}$ in terms of the $p_m(\nu_1, \cdots, \nu_k)$'s, as follows:

$$\begin{aligned}\binom{n-1}{m-1} p_{1\cdots n} &= \sum p_m(\nu_1, \cdots, \nu_m) - \sum p_m(\nu_1, \cdots, \nu_{m+1}) + \cdots \\ &\qquad + (-1)^{n-m} p_m(1, \cdots, n) \\ &= \sum_{i=0}^{n-m} (-1)^i \sum p_m(\nu_1, \cdots, \nu_{m+i}).\end{aligned} \tag{8}$$

This of course implies the following more general formula for $p_{\alpha_1 \cdots \alpha_r}$,

$$\binom{r-1}{m-1} p_{\alpha_1 \cdots \alpha_r} = \sum_{i=0}^{r-m} (-1)^i \sum p_m(\nu_1, \cdots, \nu_{m+i})$$

where $(\alpha_1, \cdots, \alpha_r)$ is a combination of the integers $(1, \cdots, n)$ and where the second summation extends to all the $\binom{r}{m+i}$-combinations of $(\alpha_1, \cdots, \alpha_r)$. Since it is known[3] that we can express other functions such as S_r, $p_{[\mu_1 \cdots \mu_r]}$ in terms of the $p_{\mu_1 \cdots \mu_r}$'s, we can also express them in terms of the $p_m(\nu_1, \cdots, \nu_k)$'s, provided $r \geqq m$.

Finally, for the case $m = 1$, we give in Theorem 6 an explicit formula for $p_{[1 \cdots r]}$ in terms of the $p_1(\nu_1, \cdots, \nu_k)$'s, as shown in (9),

$$\begin{aligned}p_{[1\cdots r]} &= - p_1(r+1, \cdots, n) + \sum_{\nu_1} p_1(\nu_1, r+1, \cdots, n) \\ &\qquad - \sum_{\nu_1, \nu_2} p_1(\nu_1, \nu_2, r+1, \cdots, n) + \cdots \\ &\qquad + (-1)^{r-1} \sum p_1(1, \cdots, r, r+1, \cdots, n), \\ &= \sum_{i=1}^{r} (-1)^{i-1} \sum_{(\nu_1, \cdots, \nu_i)} p_1(\nu_1, \cdots, \nu_i, r+1, \cdots, n),\end{aligned} \tag{9}$$

where $(\nu_1, \cdots, \nu_i)$ runs through all the $\binom{r}{i}$-combinations from $(1, \cdots, r)$. This of course implies the following more general formula:

$$p_{[\alpha_1 \cdots \alpha_r]} = \sum_{i=1}^{r} (-1)^{i-1} \sum_{(\nu_1, \cdots, \nu_i)} p_1(\nu_1, \cdots, \nu_i, \alpha_{r+1}, \cdots, \alpha_n),$$

[3, 4] Fréchet, "Condition d'existence de systemes d'événements associés a certaines probabilités," *Jour. de Math.*, (1940), p. 51-62.

where $(\alpha_1, \cdots, \alpha_r, \cdots \alpha_n)$ is a permutation of $(1, \cdots, n)$ and where $(\nu_1, \cdots, \nu_i)$ runs through all the $\binom{r}{i}$-combinations out of $(\alpha_1, \cdots, \alpha_r)$. From Theorem 6 and two lemmas we deduce a condition of existence of systems of events associated with the probabilities $p_1(\nu_1, \cdots, \nu_m)$. The author has not been able to obtain similar elegant results for the general m. Probably they do not exist.

2. Generalization of Poincaré's formula; Generalization and sharpening of Boole's inequality.

THEOREM 1:

$$p_m(1, \cdots, n) = \sum p_{\nu_1 \cdots \nu_m} - \binom{m}{1} \sum p_{\nu_1 \cdots \nu_{m+1}} + \binom{m+1}{2} \sum p_{\nu_1 \cdots \nu_{m+2}} - \cdots + (-1)^{n-m} \binom{n-1}{n-m} p_{1 \cdots n}. \tag{3}$$

PROOF: We have

$$p_m(1, \cdots, n) = \sum_{b=0}^{n-m} \sum p_{[\mu_1 \cdots \mu_{m+b}]}, \tag{10}$$

where the second summation extends, for a fixed b, to all the $\binom{n}{m+b}$-combinations of $(1, \cdots, n)$. Further we have

$$p_{\nu_1 \cdots \nu_{m+c}} = \sum_{d=0}^{n-m-c} \sum p_{[\nu_1 \cdots \nu_{m+c} \cdots \nu_{m+c+d}]} \tag{11}$$

where the second summation extends, for a fixed d, to all the $\binom{n-m-c}{d}$-combinations of $(1, \cdots, n) - (\nu_1, \cdots, \nu_{m+c})$. The formulas (10) and (11) are evident by observing that the probabilities in the summations are all additive. Now we count the number of times a fixed $p_{[\mu_1 \cdots \mu_{m+b}]}$ appears in (3). By (11) this is equal to the sum

$$\binom{m+b}{m} - \binom{m}{1}\binom{m+b}{m+1} + \binom{m+1}{2}\binom{m+b}{m+2} - \cdots + (-1)^{n-m}\binom{n-1}{n-m}\binom{m+b}{m+b} = 1,$$

since this number is the coefficient of $(-1)^m x^m$ in the expansion of

$$(1-x)^{m+b}\left(1 - \frac{1}{x}\right)^{-m} = (-1)^{-m} x^m (1-x)^b.$$

Thus by (10) we have (3).

THEOREM 2: *For $2l \leqq n - m$ and $2l \leqq n - m$ respectively, we have*

$$(6) \quad \sum_{i=0}^{2l+1} (-1)^i \binom{m+i-1}{i} S_{m+i} \leqq p_m(1, \cdots, n) \leqq \sum_{i=0}^{2l} (-1)^i \binom{m+i-1}{i} S_{m+i}.$$

PROOF: By the reasoning in the previous proof, it is sufficient (in fact also necessary) to show that

$$\sum_{i=0}^{2l} \binom{m-1+i}{i}\binom{m+b}{m+i} \geqq 1, \qquad \sum_{i=0}^{2l+1} \binom{m-1+i}{i}\binom{m+b}{m+i} < 1.$$

Since

$$\binom{m-1+i}{i}\binom{m+b}{m+i} = \frac{(m+b)!}{(m-1)!\,b!}\binom{b}{i}\frac{1}{m+i}$$

is an integer, it is sufficient to show that

$$(12) \qquad \sum_{i=0}^{2l} (-1)^i \binom{b}{i} \frac{1}{m+i} > 0, \qquad \sum_{i=0}^{2l+1} (-1)^i \binom{b}{i} \frac{1}{m+i} \leqq 0.$$

Suppose $b > 0$ is even. For $i \leqq b/2 - 1$, we have $\frac{b-i}{i+1} > 1$ so that $\frac{b-i}{i+1} \geqq \frac{i+2}{i+1}$. Also $\frac{m+i}{m+i+1} \geqq \frac{i+1}{i+2}$ for $m \geqq 1$. Hence

$$\binom{b}{i+1}\frac{1}{m+i+1} = \frac{b-i}{i+1}\,\frac{m+i}{m+i+1}\binom{b}{i}\frac{1}{m+i}$$
$$\geqq \frac{i+2}{i+1}\,\frac{i+1}{i+2}\binom{b}{i}\frac{1}{m+i} = \binom{b}{i}\frac{1}{m+i}.$$

For $i \geqq b/2$ we have $\frac{b-i}{i+1} < 1$ so that $\frac{b-i}{i+1}\,\frac{m+i}{m+i+1} < 1$ and

$$\binom{b}{i+1}\frac{1}{m+i+1} < \binom{b}{i}\frac{1}{m+i}.$$

Thus the absolute values of the terms of the alternating series

$$\sum_{i=0}^{b} (-1)^i \binom{b}{i}\frac{1}{m+i} = \frac{b!}{(m+b)!(m-1)!}$$

are monotone increasing as long as $i \leqq \frac{b}{2} - 1$, reaching maximum at $i = \frac{b}{2}$ and then become monotone decreasing.

Therefore (12) evidently holds for $2l \leqq b/2$ and $2l + 1 \leqq b/2$ respectively. For $t \geqq \frac{b}{2} + 1$ we write

$$\sum_{i=0}^{t} (-1)^i \binom{b}{i}\frac{1}{m+i} = \frac{b!}{(m+b)!(m-1)!} - \sum_{i=t+1}^{b} (-1)^i \binom{b}{i}\frac{1}{m+i}$$
$$= \frac{b!}{(m+b)!(m-1)!} - \sum_{j=0}^{b-t-1} (-1)^j \binom{b}{j}\frac{1}{m+b-j}.$$

From the above and the fact that $\dfrac{b!}{(m+b)!(m-1)!} \leqq \dfrac{1}{m+b}$ we see that the righthand side is an alternating series whose terms are non-decreasing in absolute values. Hence (12) is true.

If b is odd, the case is similar.

3. Generalization of Fréchet's inequalities and related inequalities. Before proving our remaining theorems, we shall give a more detailed account of the general method which will be used. In the foregoing work we have already given two different expressions for the function $p_m(1, \cdots, n)$, namely, formulas (3) and (10), but they are not convenient for our later purposes. Formula (3) is inconvenient because it is not additive and because the $p_{\nu_1 \cdots \nu_i}$'s are related in magnitudes; while formula (10) has gone so far in the separation of the additive constituents that its application raises algebraical difficulties. Let us therefore take an intermediate course.

Let each $\binom{n}{m}$-combination $(\nu_1, \cdots, \nu_m)$ out of $(1, \cdots, n)$ be written so that $\nu_1 < \nu_2 < \cdots < \nu_m$. Then we arrange them in an ordered sequence in the following way: the combination $(\nu_1, \cdots, \nu_m)$ is to precede the combination $(\mu_1, \cdots, \mu_m)$ if, for the first $\nu_i \neq \mu_i$, we have $\nu_i > \mu_i$. After such an arrangement we symbolically denote these combinations by

$$\mathrm{I}, \mathrm{II}, \cdots, \left[\binom{n}{m}\right].$$

Further, all the $\binom{k}{m}$-combinations out of $(\nu_1, \cdots, \nu_k)$ where the latter is a combination out of $(1, \cdots, n)$ are arranged in the order in which they appear in the sequence just written. For example, all the $\binom{4}{2}$-combinations out of $(1, 2, 3, 4)$ are ordered thus:

$$(12) \quad (13) \quad (14) \quad (23) \quad (24) \quad (34).$$

Let U denote a typical combination $(\mu_1, \cdots, \mu_m)$. By E_U we mean the combination of events $E_{\mu_1}, \cdots, E_{\mu_m}$ so that $p_U = p_{\mu_1 \cdots \mu_m}$. In general, let the combinations $U_1, \cdots, U_{b-1}, U_b$ be given, then $p_{U_1' \cdots U_{b-1}' U_b}$ denotes the probability of the non-occurrence of $U_1, \cdots, U_{b-1}$ and the occurrence of U_b.

Now let $\mathrm{I}, \mathrm{II}, \cdots, \left[\binom{k}{m} - 1\right] = Y, \left[\binom{k}{m}\right] = Z$ denote all the $\binom{k}{m}$-combinations out of $(\nu_1, \cdots, \nu_k)$ in their assigned order. We have

$$p_m(\nu_1, \cdots, \nu_k) = p_{\mathrm{I}} + p_{\mathrm{I}'\mathrm{II}} + p_{\mathrm{I}'\mathrm{II}'\mathrm{III}} + \cdots + p_{\mathrm{I}' \cdots Y' Z}. \tag{13}$$

This fundamental formula is evident. Of course it is possible to identify the p's on the right-hand side with the ordinary $p_{\nu_1' \cdots \nu_i}$'s, but we shall refrain from so doing and be content with the following example:

$$p_2(1, 2, 3, 4) = p_{12} + p_{12'3} + p_{12'3'4} + p_{1'23} + p_{1'23'4} + p_{1'2'34}.$$

THEOREM 3. *For $k = 1, \cdots, n - 1$ and $1 \leqq m \leqq k$ we have*

$$\binom{n-m}{k-m} \Sigma p_m(\nu_1, \cdots, \nu_{k+1}) \leqq \binom{n-m}{k+1-m} \Sigma p_m(\nu_1, \cdots, \nu_k).$$

PROOF. Substitute (13) and a similar formula for $k + 1$ into the two sides respectively. After this substitution we observe that the number of terms is the same on both sides, since

$$\binom{n-m}{k-m}\binom{n}{k+1}\binom{k+1}{m} = \binom{n-m}{k+1-m}\binom{n}{k}\binom{k}{m}.$$

Also, the number of terms with a given $U = (\mu_1, \cdots, \mu_m)$ unaccented is the same, since

$$\binom{n-m}{k-m}\binom{n-m}{k+1-m} = \binom{n-m}{k+1-m}\binom{n-m}{k-m}.$$

Let the sum of all the terms with U unaccented in the two summations be denoted by $\sigma_{k+1} = \sigma_{k+1}(\mu_1, \cdots, \mu_m)$ and $\sigma_k = \sigma_k(\mu_1, \cdots, \mu_m)$ respectively. It is sufficient to prove that

$$\binom{n-m}{k-m} \sigma_{k+1} \leqq \binom{n-m}{k+1-m} \sigma_k, \tag{14}$$

for any U. σ_k contains $\binom{n-m}{k-m}$ terms each of the form $p_{\nu_1' \cdots \nu_l' \mu_1 \cdots \mu_m}$ where $0 \leqq l \leqq \mu_m - m$ and where $(\nu_1, \cdots, \nu_l, \mu_1, \cdots, \mu_m)$ is a $\binom{\mu_m}{m+l}$-combination out of $(1, \cdots, \mu_m)$. For fixed $(\mu_1, \cdots, \mu_m)$ and a fixed l but varying λ's, σ_k contains $\binom{n-\mu_m}{k-m-l}$ terms of the form $p_{\nu_1' \cdots \nu_l' \mu_1 \cdots \mu_m}$, with exactly l accented subscripts. Let the sum of all such terms be denoted by $\sigma_k^{(l)}$. Evidently $\sigma_k^{(l)}$ has $\binom{\mu_m - m}{l}$ terms. As a check we have

$$\binom{n-\mu_m}{k-m}\binom{\mu_m - m}{0} + \binom{n-\mu_m}{k-m-1}\binom{\mu_m-m}{1} + \cdots + \binom{n-\mu_m}{k-\mu_m}\binom{\mu_m-m}{\mu_m-m} = \binom{n-m}{k-m},$$

which is the total number of terms in σ_k.

We decompose these p's partially, as follows:

$$p_{\nu_1' \cdots \nu_l' \mu_1 \cdots \mu_m} = \sum_{b=0}^{\mu_m - m - l} \sum_{\mu_{m+1}, \cdots, \mu_{m+b}} p_{\nu_1' \cdots \nu_{l+c}' \mu_1 \cdots \mu_{m+b}},$$

where $(\nu_1, \cdots, \nu_{l+c}, \mu_1, \cdots, \mu_{m+b})$ is a permutation of $(1, \cdots, \mu_m)$ and where the second summation extends, for a fixed b, to all the $\binom{\mu_m - m - l}{b}$-combinations out of $(1, \cdots, \mu_m) - (\nu_1, \cdots, \nu_l, \mu_1, \cdots \mu_m)$.

Now consider a given

$$p_{\rho_1\cdots\rho_t\lambda_1\cdots\lambda_s\mu_1\cdots\mu_m}$$

where $0 \leqq t \leqq \mu_m - m$ and $(\rho_1 \cdots \rho_t\lambda_1 \cdots \lambda_s\mu_1 \cdots \mu_m)$ is a permutation of $(1, \cdots, \mu_m)$. It appears $\binom{t}{l}$ times in $\sigma_k^{(l)}$. Hence it appears

$$\binom{n-\mu_m}{k-m}\binom{t}{0} + \binom{n-\mu_m}{k-m-1}\binom{t}{1} + \cdots + \binom{n-\mu_m}{k-m-t}\binom{t}{t} = \binom{n-\mu_m+t}{k-m}$$

times in σ_k.

Therefore to prove (14) it is sufficient to prove that

$$\binom{n-m}{k-m}\binom{n-\mu_m+t}{k+1-m} \leqq \binom{n-m}{k+1-m}\binom{n-\mu_m+t}{k-m}.$$

By an easy reduction we have

$$(n - \mu_m + t - k + m) \leqq n - k$$

or

$$-\mu_m + t + m \leqq 0;$$

since $t \leqq \mu_m - m$ this is obvious.

THEOREM 4: *For* $2 \leqq k \leqq n - 1$ *and* $1 \leqq m \leqq k$ *we have*

$$(5) \qquad \frac{\Sigma p_m(\nu_1, \cdots, \nu_k)}{\binom{n-m}{k-m}} \leqq \frac{1}{2}\,\frac{\Sigma p_m(\nu_1, \cdots, \nu_{k-1})}{\binom{n-m}{k-1-m}} + \frac{1}{2}\,\frac{\Sigma p_m(\nu_1, \cdots, \nu_{k+1})}{\binom{n-m}{k+1-m}}.$$

PROOF: By the reasoning in the previous proof, it is sufficient to show that

$$\begin{aligned}2\binom{n-m}{k-1-m}&\binom{n-m}{k+1-m}\binom{n-\mu_m+t}{k-m}\\ &\leqq \binom{n-m}{k-m}\binom{n-m}{k+1-m}\binom{n-\mu_m+t}{k-1-m}\\ &\qquad + \binom{n-m}{k-m}\binom{n-m}{k-1-m}\binom{n-\mu_m+t}{k+1-m},\end{aligned}$$

for $0 \leqq t \leqq \mu_m - m$. By an easy reduction this is equivalent to

$$\begin{aligned}2(n-k)(n-\mu_m+t-k+m+1) &\leqq (n-k+1)(n-k)\\ &\quad + (n-\mu_m+t-k+m+1)(n-\mu_m+t-k+m)\end{aligned}$$

or

$$(n - \mu_m + t - k + m + 1)(\mu_m - t - m) \leqq (n - k)(\mu_m - t - m).$$

For $t = \mu_m - m$ we have equality, otherwise we have

$$-\mu_m + t + m + 1 \leqq 0.$$

We can deduce Theorem 3 from Theorem 4 and the following result (a case of generalized Gumbel inequalities):

$$\binom{n-1}{n-2} p_m(1, \cdots, n) \leqq \Sigma p_m(\nu_1, \cdots, \nu_{n-1}). \tag{15}$$

PROOF OF (15): Substitute from (13). Consider the p's with U unaccented. The number of such terms is the same on both sides. But on the left-hand side they are all the same $p_{1'2'\cdots(U-1)'U}$, while those on the right-hand side, being of the form $p_{U_1'\cdots U_\lambda' U}$ where $0 \leqq \lambda \leqq U - 1$ and $(U_1, \cdots, U_\lambda)$ is a combination out of $(1, \cdots, U - 1)$, are greater than or equal to it. Hence the result.

4. The $p_{\alpha_1\cdots\alpha_i}$'s in terms of the $p_m(\nu_1, \cdots, \nu_k)$'s and the $p_{[\alpha_1\cdots\alpha_i]}$'s in terms of the $p_1(\nu_1, \cdots, \nu_k)$'s.

THEOREM 5: *For* $1 \leqq m \leqq n$ *we have*

$$\begin{aligned}\binom{n-1}{m-1} p_{1\cdots n} &= \sum p_m(\nu_1, \cdots, \nu_m) - \sum p_m(\nu_1, \cdots, \nu_{m+1}) + \cdots \\ &\qquad + (-1)^{n-m} p_m(1, \cdots, n) \\ &= \sum_{i=0}^{n-m} (-1)^i \sum_{\nu_1, \cdots \nu_{m+i}} p_m(\nu_1, \cdots, \nu_{m+i}).\end{aligned} \tag{8}$$

PROOF: As in the proof of Theorem 3, consider $\sigma_k(\mu_1, \cdots, \mu_m)$. Here $m \leqq k \leqq n$. Since a given

$$p_{\rho_1'\cdots\rho_t'\lambda_1\cdots\lambda_s\mu_1\cdots\mu_m}, \tag{16}$$

appears $\binom{n-\mu_m+t}{k-m}$ times in σ_k, it appears

$$\begin{aligned}\sum_{k=m}^{n} (-1)^{k-m} \binom{n-\mu_m+t}{k-m} &= \sum_{j=0}^{n-m} (-1)^j \binom{n-\mu_m+t}{j} \\ &= \sum_{j=0}^{n-\mu_m+t} (-1)^j \binom{n-\mu_m+t}{j} = \begin{cases} 0, & \text{if } n-\mu_m+t \geqq 1, \\ 1, & \text{if } n-\mu_m+t = 0.\end{cases}\end{aligned}$$

times on the right hand side of (8). Hence for fixed $(\mu_1, \cdots, \mu_m)$, the only p's of the form (16) which actually appears are those with $t = \mu_m - n$. But $\mu_m \leqq n$, thus $t = 0$, $\mu_m = n$, and $(\lambda_1, \cdots \lambda_s, \mu_1, \cdots, \mu_m)$ is a permutation of $(1, \cdots, n)$. The term in question is therefore $p_{1\cdots n}$. Since the number of $\binom{n}{m}$-combinations of $(1, \cdots, n)$ with $\mu_m = n$ is $\binom{n-1}{m-1}$, we have the theorem.

THEOREM 6: *For* $1 \leqq r \leqq n - 1$, *we have*

$$\begin{aligned}p_{[1\cdots r]} &= -\,p_1(r+1, \cdots, n) + \sum_{\nu_1} p_1(\nu_1, r+1, \cdots, n) \\ &\quad - \sum_{\nu_1, \nu_2} p(\nu_1, \nu_2, r+1, \cdots, n) + \cdots + (-1)^{r-1} \sum p_1(1, \cdots, n) \\ &= \sum_{i=1}^{r} (-1)^{i-1} \sum_{\nu_1, \cdots, \nu_i} p_1(\nu_1, \cdots, \nu_i, r+1, \cdots, n),\end{aligned} \tag{9}$$

where $(\nu_1, \cdots, \nu_i)$ *runs through all the* $\binom{r}{i}$*-combinations out of* $(1, \cdots, r)$.

PROOF: We rewrite (14) for the special case $m = 1$,

$$p_1(\mu_1, \cdots, \mu_k) = p_{\mu_1} + p_{\mu_1'\mu_2} + \cdots + p_{\mu_1'\cdots\mu_{k-1}'\mu_k}, \tag{17}$$

where $\mu_1 < \mu_2 < \cdots < \mu_k$. Substitute into the right hand side of (9). After the substitution let the sum of all those p's with μ unaccented be denoted by σ_μ. The terms in σ_μ are of the form $p_{\mu_1'\cdots\mu_{s-1}'\mu}$ where $1 \leqq s \leqq \mu$ and $(\mu_1, \cdots, \mu_{s-1})$ is a combination out of $(1, \cdots, \mu - 1)$.

First consider a fixed $\mu \leqq r$. For a fixed $p_{\mu_1'\cdots\mu_{s-1}'\mu}$ we count the number of times it appears in σ_μ, that is, on the right hand side of (9). This is evidently equal to

$$\sum_{j=s}^{r} (-1)^j \binom{r-\mu}{j-s} = \sum_{j=s}^{r-\mu+s} (-1)^j \binom{r-\mu}{j-s} = \begin{matrix} 0, & \text{if } r - \mu \geqq 1, \\ 1, & \text{if } r - \mu = 0. \end{matrix}$$

Thus the only terms that actually appear are those with $\mu = r$; and each of such terms $p_{\mu_1'\cdots\mu_{s-1}'r}$ appears exactly once with the sign $(-1)^s$. Hence their total contribution is

$$p_r - \sum_{\nu_1} p_{\nu_1'r} + \sum_{\nu_1,\nu_2} p_{\nu_1'\nu_2'r} - \cdots + (-1)^{r-1} p_{1'\cdots(r-1)'r} = p_{1\cdots r}, \tag{18}$$

by an easy modification of Poincaré's formula.

Next consider a fixed $\mu \geqq r + 1$. Every term with μ unaccented in σ_μ is of the form (with the usual convention for $\mu = r + 1$) $p_{\mu_1'\cdots\mu_s'(r+1)'\cdots(\mu-1)'\mu}$, where $(\mu_1, \cdots, \mu_s)$ is a combination out of $(1, \cdots, r)$; and it appears exactly once with the sign $(-1)^s$. Their total contribution is therefore

$$-p_{(r+1)'\cdots(\mu-1)'\mu} + \sum_{\nu_1} p_{\nu_1'(r+1)'\cdots(\mu-1)'\mu} - \sum_{\nu_1,\nu_2} p_{\nu_1'\nu_2'(r+1)'\cdots(\mu-1)'\mu} + \cdots$$
$$+ (-1)^{r-1} p_{1'\cdots(\mu+1)'\mu} = -p_{1\cdots r(r+1)'\cdots(\mu-1)'\mu},$$

by another application of Poincaré's formula. Summing up for $\mu = r + 1, \cdots, n$, we obtain

$$-(p_{1\cdots r(r+1)} + p_{1\cdots r(r+1)'(r+2)} + \cdots + p_{1\cdots r(r+1)'\cdots(n-1)'n}). \tag{19}$$

Adding (18) and (19), we obtain as the sum of the right-hand side of (9)

$$p_{1\cdots r} - (p_{1\cdots r(r+1)} + p_{1\cdots r(r+1)'(r+2)} + \cdots + p_{1\cdots r(r+1)'\cdots(n-1)'n})$$
$$= p_{1\cdots r(r+1)'(r+2)'\cdots n'} = p_{[1\cdots r]}$$

by an easy modification of (17).

5. A condition for existence of systems of events associated with the probabilities $p_1(\nu_1, \cdots, \nu_k)$.

LEMMA 1: *Let any* $2^n - 1$ *quantities* $q(\alpha_1, \cdots, \alpha_k)$ *be given, where* $k =$

$1, \cdots, n$, *and for a fixed* k, $(\alpha_1, \cdots, \alpha_k)$ *runs through all the* $\binom{n}{k}$*-combinations out of* $(1, \cdots, n)$. *Let the quantities* $Q(\alpha_1, \cdots, \alpha_k)$ *be formed as follows:*

$$Q(0) = 1 - q(1, \cdots, n),$$

$$Q(\alpha_1, \cdots, \alpha_k) = -q(\alpha_{k+1}, \cdots, \alpha_n) + \sum_{\nu_1} q(\nu_1, \alpha_{k+1}, \cdots, \alpha_n)$$
$$- \sum_{\nu_1, \nu_2} q(\nu_1, \nu_2, \alpha_{k+1}, \cdots, \alpha_n) + \cdots + (-1)^{k-1} q(1, \cdots, n),$$

where $(\nu_1, \cdots, \nu_i)$ runs through all the $\binom{k}{i}$-combinations out of $(1, \cdots, n) - (\alpha_{k+1}, \cdots, \alpha_n)$. *Then the sum of all these* Q*'s is equal to* 1.

PROOF: Add all these Q's and count the number of times a fixed $q(\mu_1, \cdots, \mu_k)$ appears in the sum. For $1 \leqq k \leqq n$ this number is equal to

$$-1 + \binom{k}{1} - \binom{k}{2} + \cdots + (-1)^{k-1}\binom{k}{k} = 0.$$

Hence we have the lemma.

LEMMA 2: (Fréchet) *Given* 2^n *quantities* $Q_{[\alpha_1 \cdots \alpha_r]}$ *where* $(\alpha_1, \cdots, \alpha_r)$ *runs through all combinations out of* $(1, \cdots, n)$ *including the empty one. The necessary and sufficient condition that there exist systems of events* $E_1, \cdots, E_n$ *for which*

$$p_{[\alpha_1 \cdots \alpha_r]} = Q_{[\alpha_1 \cdots \alpha_r]}$$

(where $p_{[0]}$ denotes the probability for the non-occurrence of $E_1, \cdots, E_n$) *is that each* $Q \geqq 0$ *and that their sum is equal to* 1.

PROOF: Since the probabilities $p_{[\alpha_1 \cdots \alpha_r]}$ are independent, i.e., unrelated in magnitudes except that their sum is equal to 1, the lemma is evident.

THEOREM 7: *Given* $2^n - 1$ *quantities* $q(\alpha_1, \cdots, \alpha_k)$ *as in* Lemma 1, *the necessary and sufficient condition that there exist systems of events* $E_1, \cdots, E_n$ *for which*

$$p_1(\alpha_1, \cdots, \alpha_k) = q(\alpha_1, \cdots, \alpha_k)$$

is that for any combination $(\alpha_{r+1}, \cdots, \alpha_n)$, $1 \leqq r \leqq n - 1$, *out of* $(1, \cdots, n)$ *we have*

$$-q(\alpha_{r+1}, \cdots, \alpha_n) + \sum_{\nu_1} q(\alpha_{\nu_1}, \alpha_{r+1}, \cdots, \alpha_n) - \sum_{\nu_1, \nu_2} q(\alpha_{\nu_1}, \alpha_{\nu_2}, \alpha_{r+1}, \cdots, \alpha_n)$$
$$+ \cdots + (-1)^{r-1} q(1, \cdots, n) \geqq 0,$$

and thus

$$1 - q(1, \cdots, n) \geqq 0.$$

PROOF: The condition is necessary by Theorem 6. It is sufficient by Lemma 1, 2 and an obvious formula expressing $p_1(\alpha_1, \cdots, \alpha_r)$ in terms of the $p_{[\nu_1 \cdots \nu_i]}$'s.

Reprinted from *Ann. Math. Stat.* 13 (1942), pp. 338–349
with permission from Institute of Mathematical Statistics.

ON MUTUALLY FAVORABLE EVENTS

BY KAI-LAI CHUNG

Tsing Hua University, Kunming, China

Introduction. For a set of arbitrary events, E. J. Gumbel, M. Fréchet and the author[1] have recently obtained inequalities between sums of certain probability functions. One of the results of the author is the following:

Let $E_1, \cdots, E_n$ be n arbitrary events and let $p_m(\nu_1, \cdots, \nu_k)$ denote the probability of the occurrence of at least m events out of the k events $E_{\nu_1}, \cdots, E_{\nu_k}$. Then, for $k = 1, \cdots, n-1$ and $1 \leqq m \leqq k$ we have

$$\binom{n-m}{k-m}\Sigma p_m(\nu_1, \cdots, \nu_{k+1}) \leqq \binom{n-m}{k-m+1}\Sigma p_m(\nu_1, \cdots, \nu_k),$$

where the summations extend respectively to all combinations of $k+1$ and k indices out of the n indices $1, \cdots, n$.

In course of proof of the above inequalities it appears that similar inequalities between products instead of sums can be obtained under certain assumptions regarding the nature of interdependence of the events. We shall first study the nature of such assumptions, and then proceed to the proof of the said inequalities (Theorems 1 and 2). It may be noted that the inductive method used here serves equally well for the proof of the inequalities cited above, though somewhat longer, but apparently our former method is not applicable here.

That events satisfying our assumptions actually exist, is shown by an application to the elementary theory of numbers. The author feels incompetent to discuss other possible fields of application.

1. Let a set of events be given

$$E_1, E_2, \cdots, E_n, \cdots$$

and let E_i' denote the event non-E_i. Let $p(i)$ denote the probability of the occurrence of E_i, $p(i')$ that of the occurrence of E_i'. For convenience we assume that for any i $p_i(1 - p_i) \neq 0$; events with the exceptional probabilities 0 or 1 may evidently be left out of account.

Let $p(\nu_1 \cdots \nu_k)$ denote the probability of the occurrence of the conjunction $E_{\nu_1} \cdots E_{\nu_k}$ and let $p(\mu_1 \cdots \mu_h, \nu_1 \cdots \nu_k)$ denote the probability of the occurrence of $E_{\nu_1} \cdots E_{\nu_k}$, on the hypothesis that $E_{\mu_1} \cdots E_{\mu_h}$ have occurred. The μ's or ν's may be accented.

DEFINITION 1: *If $p(\nu_1, \nu_2) > p(\nu_2)$, we say that the occurrence of the event E_{ν_1} is favorable to the occurrence of the event E_{ν_2}, or simply that E_{ν_1} is favorable to E_{ν_2}.*

[1] "On the probability of the occurrence of at least m events among n arbitrary events," *Annals of Math. Stat.* Vol. 12 (1941), pp. 328–338.

If $p(\nu_1, \nu_2) = p(\nu_2)$, we say that E_{ν_1} is indifferent to E_{ν_2}. If $p(\nu_1, \nu_2) < p(\nu_2)$, we say that E_{ν_1} is unfavorable to E_{ν_2}.

Thus the relations "favorableness," "indifference," and "unfavorableness" are mutually exclusive and together exhaustive. We state the following immediate consequences:

(i) Reflexity: An event is favorable to itself; in fact, $p(\nu, \nu) = 1 > p(\nu)$.

(ii) Symmetry: If E_1 is favorable (indifferent, unfavorable) to E_2, then E_2 is favorable (indifferent, unfavorable) to E_1. In fact, we have

$$p(1)p(1, 2) = p(12) = p(2)p(2, 1),$$

$$\frac{p(1, 2)}{p(2)} = \frac{p(2, 1)}{p(1)}.$$

Thus $p(1, 2) \gtreqless p(2)$ is equivalent to $p(2, 1) \gtreqless p(1)$.

In particular, if E_1 is indifferent to E_2, then so is E_2 to E_1. They are then usually said to be independent of each other.

(iii) If E_1 is favorable (indifferent, unfavorable) to E_2, then E_1' is unfavorable (indifferent, favorable) to E_2. For, we have

$$p(1)p(1, 2) + p(1')p(1', 2) = p(12) + p(1'2) = p(2),$$

whence

$$p(1')p(1', 2) = p(2) - p(1)p(1, 2).$$

On the other hand,

$$p(1')p(2) = [1 - p(1)]p(2) = p(2) - p(1)p(2).$$

Since by assumption $p(1')p(2) \neq 0$, we have

$$\frac{p(1', 2)}{p(2)} = \frac{p(2) - p(1)p(1, 2)}{p(2) - p(1)p(2)}.$$

Thus

$$p(1', 2) \gtreqless p(2) \text{ according as } p(1, 2) \lesseqgtr p(2).$$

For the sake of brevity we introduce the following symbolic notation:

$$E_1/E_2 = \begin{cases} 1, \text{ if } E_1 \text{ is favorable to } E_2 \\ 0, \text{ if } E_1 \text{ is indifferent to } E_2 \\ -1, \text{ if } E_1 \text{ is unfavorable to } E_2. \end{cases}$$

Then by (ii) and (iii) we have

$$E_1/E_2 = E_2/E_1,$$

$$E_1'/E_2 = E_2/E_1' = E_1/E_2' = E_2'/E_1 = -(E_1/E_2),$$

$$E_1'/E_2' = E_2'/E_1' = E_1/E_2,$$

analogous to the rules of signs in the multiplication of integers.

(iv) Non-transitivity: If E_1 is favorable to E_2, and E_2 is favorable to E_3, it does not necessarily follow that E_1 is favorable to E_3; in fact, it may happen that E_1 is unfavorable to E_3. For instance, imagine 11 identical balls in a bag marked respectively with the numbers

$$-11, -10, -3, -2, -1, 2, 4, 6, 11, 13, 16.$$

Let a ball be drawn at random. Let

E_1 = (the event of the number on the ball being positive)

E_2 = (the event of the number on the ball being even)

E_3 = (the event of the number on the ball being of 1 digit)

We have

$$p(1, 2) = \tfrac{4}{6} > \tfrac{6}{11} = p(2),$$

$$p(2, 3) = \tfrac{4}{6} > \tfrac{6}{11} = p(3),$$

$$p(1, 3) = \tfrac{1}{2} < \tfrac{6}{11} = p(3).$$

(v) It may happen that $E_1/E_3 = 1$, $E_2/E_3 = 1$, but $E_1E_2/E_3 = -1$. In the example above,

$$p(2, 1) = \tfrac{4}{6} > \tfrac{6}{11} = p(1),$$

$$p(3', 1) = \tfrac{3}{5} > \tfrac{6}{11} = p(1),$$

$$p(23', 1) = \tfrac{1}{2} < \tfrac{6}{11} = p(1).$$

(vi) It may happen that $E_1/E_2 = 1$, $E_1/E_3 = 1$, but $E_1/E_2E_3 = -1$. Example:

$$p(1, 2) = \tfrac{4}{6} > \tfrac{6}{11} = p(2),$$

$$p(1, 3') = \tfrac{1}{2} > \tfrac{5}{11} = p(3'),$$

$$p(1, 23') = \tfrac{1}{6} < \tfrac{2}{11} = p(23').$$

(vii) It may happen that $E_1/E_3 = 1$, $E_2/E_3 = 1$, but the disjunction $(E_1 + E_2)/E_3 = -1$. For, by (v) we know that there exist events E_1', E_2', E_3' such that

$$E_1'/E_3' = 1, \qquad E_2'/E_3' = 1, \qquad E_1'E_2'/E_3' = -1.$$

Hence by (iii) there exist events E_1, E_2, E_3 such that

$$E_1/E_3 = 1, \qquad E_2/E_3 = 1, \qquad (E_1'E_2')'/E_3 = -1.$$

But $(E_1'E_2')' = E_1 + E_2$. Thus the last relation is $(E_1 + E_2)/E_3 = -1$.

(viii) It may happen that $E_1/E_2 = 1$, $E_1/E_3 = 1$, but $E_1/(E_2 + E_3) = -1$. This follows from (vi) as (vii) follows from (v).

After all these negative results in (iv)–(viii), we see that we cannot expect to go far without making stronger assumptions regarding the nature of inter-

dependence between the events in the set. Firstly, in view of (iv), we shall restrict ourselves to consideration of a set of events in which each event is favorable to every other. Secondly, in view of (v), we shall only consider the case where the "favorableness," as defined above, shall be cumulative in its effect, that is to say, the more events favorable to a given event have been known to occur, the more probable this given event shall be esteemed. We formulate these two conditions in mathematical terms, as follows:

DEFINITION 2: *A set of events $E_1, \cdots, E_n, \cdots$ is said to be strongly mutually favorable (in the first sense) if, for every integer h and every set of distinct indices (positive integers) $\mu_1, \cdots, \mu_h$ and ν we have*

$$p(\mu_1 \cdots \mu_h, \nu) > p(\mu_1 \cdots \mu_{h-1}, \nu).$$

This definition requires that there exist no implication relation between any event and any conjunction of events in the set; in particular, that the events are all distinct. It would be more convenient to consider the relation "favorable or indifferent to." This will be done later on. The present definitions have the advantage of being logically clear cut and also that of yielding unambiguous inequalities.

From Definition 2 we deduce the following consequences:

(1) If the set $(\mu_1^*, \cdots, \mu_j^*)$ is a sub-set of $(\mu_1, \cdots, \mu_h)$, we have

$$p(\mu_1 \cdots \mu_h, \nu) > p(\mu_1^* \cdots \mu_j^*, \nu).$$

(2) For any positive integer k and any two sets $(\nu_1, \cdots, \nu_k)$ and $(\mu_1, \cdots, \mu_h)$ where all the indices are distinct, we have

$$p(\mu_1 \cdots \mu_h, \nu_1 \cdots \nu_k) > p(\mu_1 \cdots \mu_{h-1}, \nu_1 \cdots \nu_k).$$

More generally, we have as in (1),

$$p(\mu_1 \cdots \mu_h, \nu_1 \cdots \nu_k) > p(\mu_1^* \cdots \mu_j^*, \nu_1 \cdots \nu_k).$$

PROOF: We have only to prove the first inequality. For $k = 1$ this is the assumption in Definition 2. Suppose that the inequality holds for $k - 1$, we shall prove that it holds for k, too:

$$\frac{p(\mu_1 \cdots \mu_h, \nu_1 \cdots \nu_k)}{p(\mu_1 \cdots \mu_{h-1}, \nu_1 \cdots \nu_k)} = \frac{p(\mu_1 \cdots \mu_{h-1})p(\mu_1 \cdots \mu_h)p(\mu_1 \cdots \mu_h, \nu_1 \cdots \nu_k)}{p(\mu_1 \cdots \mu_h)p(\mu_1 \cdots \mu_{h-1})p(\mu_1 \cdots \mu_{h-1}, \nu_1 \cdots \nu_k)}$$

$$= \frac{p(\mu_1 \cdots \mu_{h-1})p(\mu_1 \cdots \mu_h \nu_1 \cdots \nu_k)}{p(\mu_1 \cdots \mu_h)p(\mu_1 \cdots \mu_{h-1} \nu_1 \cdots \nu_k)}$$

$$= \frac{p(\mu_1 \cdots \mu_{h-1})p(\mu_1 \cdots \mu_h)p(\mu_1 \cdots \mu_h, \nu_1)p(\mu_1 \cdots \mu_h \nu_1, \nu_2 \cdots \nu_k)}{p(\mu_1 \cdots \mu_h)p(\mu_1 \cdots \mu_{h-1})p(\mu_1 \cdots \mu_{h-1}, \nu_1)p(\mu_1 \cdots \mu_{h-1} \nu_1, \nu_2 \cdots \nu_k)}$$

$$= \frac{p(\mu_1 \cdots \mu_h, \nu_1)}{p(\mu_1 \cdots \mu_{h-1}, \nu_1)} \frac{p(\mu_1 \cdots \mu_h \nu_1, \nu_2 \cdots \nu_k)}{p(\mu_1 \cdots \mu_{h-1} \nu_1, \nu_2 \cdots \nu_k)}$$

$$> \frac{p(\mu_1 \cdots \mu_h \nu_1, \nu_2 \cdots \nu_k)}{p(\mu_1 \cdots \mu_{h-1} \nu_1, \nu_2 \cdots \nu_k)} > 1.$$

Observe that none of the denominators vanish by our original assumption and by Definition 2.

Therefore we see that when the failure in (v) is remedied by our definition, the failure in (vi) is automatically remedied too.

2. THEOREM 1: *Let $n > 1$ and let $E_1, \cdots, E_n, \cdots$ be a set of strongly mutually favorable events (in the first sense). Then we have, for $k = 1, \cdots, n-1$,*

$$\prod_{\nu_1, \cdots, \nu_{k+1}} [p(\nu_1 \cdots \nu_{k+1})]^{\binom{n-1}{k}^{-1}} > \prod_{\nu_1, \cdots, \nu_k} [p(\nu_1 \cdots \nu_k)]^{\binom{n-1}{k-1}^{-1}}$$

where the products extend respectively to all combinations of $k + 1$ and k distinct indices out of the indices $1, \cdots, n$.

PROOF. We may assume that the indices are written so that

$$1 \leqq \nu_1 < \nu_2 < \cdots < \nu_{k+1} \leqq n.$$

Taking logarithms, we have

$$\binom{n-1}{k-1} \sum_{\nu_1, \cdots, \nu_{k+1}} \log p(\nu_1 \cdots \nu_{k+1}) > \binom{n-1}{k} \sum_{\nu_1, \cdots, \nu_k} \log p(\nu_1 \cdots \nu_k).$$

Substituting from the obvious formula

$$p(\nu_1 \cdots \nu_k) = p(\nu_1)p(\nu_1, \nu_2)p(\nu_1\nu_2, \nu_3) \cdots p(\nu_1 \cdots \nu_{k-1}, \nu_k),$$

and writing $\log p(\cdots) = q(\cdots)$, the inequality becomes

$$\begin{aligned} &\binom{n-1}{k-1} \Sigma[q(\nu_1) + q(\nu_1, \nu_2) + \cdots + q(\nu_1 \cdots \nu_k, \nu_{k+1})] \\ (1) \qquad &> \binom{n-1}{k} \Sigma[q(\nu_1) + q(\nu_1, \nu_2) + \cdots + q(\nu_1 \cdots \nu_{k-1}, \nu_k)]. \end{aligned}$$

Immediately we observe that the number of terms of the form $q(\nu_1 \cdots \nu_s, \mu)(0 \leqq s \leqq \mu - 1)$ with a fixed μ after the comma in the bracket is the same on both sides, since

$$(2) \qquad \binom{n-1}{k-1}\binom{n-1}{k} = \binom{n-1}{k}\binom{n-1}{k-1}.$$

Let the sums of such q's on the left and right of (1) be $\sigma^{(1)} = \sigma^{(1)}(\mu)$ and $\sigma^{(2)} = \sigma^{(2)}(\mu)$ respectively. To prove our theorem it is sufficient to prove that $\sigma^{(1)}(\mu) \geqq \sigma^{(2)}(\mu)$ for every μ and $\sigma^{(1)}(\mu) > \sigma^{(2)}(\mu)$ for at least one μ.

Now the terms in $\sigma^{(1)}$ (or $\sigma^{(2)}$) fall into classes according to the number s of the μ_i's before the comma in the bracket. Let those terms having s μ_i's before the comma belong to the s-th class. It is evident that the number of terms of the s-th class in $\sigma^{(1)}(\mu)$ is equal to

$$\binom{n-1}{k-1}\binom{\mu-1}{s}\binom{n-\mu}{k-s}$$

for $s = 0, 1, \cdots, \mu - 1$; where we make the convention that

$$\binom{0}{0} = 1, \qquad \binom{a}{b} = 0 \quad \text{if } a < b \text{ or if } b < 0.$$

Thus for a fixed μ, when the terms in $\sigma^{(1)}(\mu)$ are classified in the above manner, its total number of terms may be written as the following sum, in which vanishing terms may occur:

$$\begin{aligned}\binom{n-1}{k-1}\binom{n-1}{k} = \binom{n-1}{k-1}\Bigg\{&\binom{\mu-1}{\mu-1}\binom{n-\mu}{k-\mu+1} \\ &+ \binom{\mu-1}{\mu-2}\binom{n-\mu}{k-\mu+2} + \cdots + \binom{\mu-1}{s}\binom{n-\mu}{k-s} \\ &+ \cdots + \binom{\mu-1}{0}\binom{n-\mu}{k}\Bigg\}.\end{aligned}$$

Similarly the total number of terms in $\sigma^{(2)}(\mu)$ may be written as the following sum:

$$\begin{aligned}\binom{n-1}{k}\binom{n-1}{k-1} = \binom{n-1}{k}\Bigg\{&\binom{\mu-1}{\mu-1}\binom{n-\mu}{k-\mu} + \binom{\mu-1}{\mu-2}\binom{n-\mu}{k-\mu+1} \\ &+ \cdots + \binom{\mu-1}{s}\binom{n-\mu}{k-s-1} + \cdots + \binom{\mu-1}{0}\binom{n-\mu}{k-1}\Bigg\}.\end{aligned}$$

Lemma 1: *For $0 \leqq s \leqq k$, we have, taking account of our conventions about the binomial coefficients,*

$$(3) \quad \binom{n-1}{k-1}\binom{n-\mu}{k-s} > \binom{n-1}{k}\binom{n-\mu}{k-s-1} \quad \textit{for} \quad s > (\mu-1)k/n;$$

$$(4) \quad \binom{n-1}{k-1}\binom{n-\mu}{k-s} \leqq \binom{n-1}{k}\binom{n-\mu}{k-s-1} \quad \textit{for} \quad s \leqq (\mu-1)k/n.$$

Proof: Suppose $s \geqq k - n + \mu$, then

$$\binom{n-1}{k-1}\binom{n-\mu}{k-s} \gtreqless \binom{n-1}{k}\binom{n-\mu}{k-s-1}$$

according as

$$\frac{k}{n-k} \gtreqless \frac{k-s}{n-\mu-k+s+1},$$

i.e. according as

$$s \gtreqless (\mu-1)k/n.$$

But, since $k < n$ and $\mu \leqq n$, we have

$$n - k - k/n + 1 > (n-k)\mu/n$$

$$(\mu-1)k/n > k - n + \mu - 1$$

so that

$$(\mu - 1)k/n + 1 \geqq k - n + \mu.$$

Therefore if $s > (\mu - 1)k/n$, then $s \geqq (\mu - 1)k/n + 1 \geqq k - n + \mu$, and (3) holds.

Again, if $k - n + \mu \leqq s \leqq (\mu - 1)k/n$, then (4) holds; while if $s < k - n + \mu$, then the left-hand side of (4) vanishes while the right-hand side is non-negative, thus (4) holds for $s \leqq (\mu - 1)k/n$. The lemma is proved.

If we put $(s = 0, 1, \cdots, k)$

$$\binom{n-1}{k-1}\binom{n-\mu}{k-s} - \binom{n-1}{k}\binom{n-\mu}{k-s-1} = d_s,$$

then by Lemma 1,

$$d_s \gtreqless 0 \quad \text{according as} \quad s \gtreqless (\mu - 1)k/n.$$

This means that although the total number of terms of the form $p(\mu_1 \cdots \mu_s, \mu)$ is the same on both sides of (1), the left-hand side is more abundant in terms with larger s while the right-hand side is more abundant in terms with smaller s. Now we have

$$q(\mu_1 \cdots \mu_i, \mu) > q(\mu_1^* \cdots \mu_j^*, \mu)$$

if $i > j$ and if $(\mu_1^* \cdots \mu_j^*)$ is a subset of $(\mu_1 \cdots \mu_i)$. Hence it is natural to suppose that the left-hand side must be greater because it is more abundant in terms of larger values. Unfortunately even if $i > j$, the last inequality is in general not true if the set $(\mu_1^* \cdots \mu_j^*)$ is not a sub-set of $(\mu_1 \cdots \mu_i)$. Therefore we cannot as yet conclude that $\sigma^{(1)} \geqq \sigma^{(2)}$.

To prove that actually we have $\sigma^{(1)} \geqq \sigma^{(2)}$, we make the following "process of compensation":

We have, by (2) and the definition of d_s, the following equality:

$$\binom{\mu-1}{0}d_0 + \binom{\mu-1}{1}d_1 + \cdots + \binom{\mu-1}{\mu-1}d_{\mu-1} = 0.$$

where $d_j = 0$ if $j > k$. Thus

$$d_s \leqq 0 \quad \text{for} \quad s \leqq k(\mu - 1)/n,$$

$$d_s \geqq 0 \quad \text{for} \quad s > k(\mu - 1)/n.$$

Hence

$$(5) \qquad \binom{\mu-1}{0}d_0 + \binom{\mu-1}{1}d_1 + \cdots + \binom{\mu-1}{\mu-1}d_{\mu-1} \leqq 0$$

for $l = 0, 1, \cdots, \mu - 1$.

For the fixed μ, let

$$\rho_l^{(1)} = \binom{n-1}{k-1}\left\{\binom{n-\mu}{k}q_\mu + \binom{n-\mu}{k-1}\sum_{\mu_1<\mu} q(\mu_1, \mu) + \cdots \right.$$
$$\left. + \binom{n-\mu}{k-l}\sum_{\mu_1<\cdots<\mu_l<\mu} q(\mu_1 \cdots \mu_l, \mu)\right\}$$

$$\rho_l^{(2)} = \binom{n-1}{k}\left\{\binom{n-\mu}{k-1}q_\mu + \binom{n-\mu}{k-2}\sum_{\mu_1<\mu} q(\mu_1, \mu) + \cdots \right.$$
$$\left. + \binom{n-\mu}{k-l-1}\sum_{\mu_1<\cdots<\mu_l<\mu} q(\mu_1 \cdots \mu_l, \mu)\right\}$$

so that

$$\rho_{\mu-1}^{(1)} = \sigma^{(1)}(\mu), \qquad \rho_{\mu-1}^{(2)} = \sigma^{(2)}(\mu).$$

For $\mu = 1$, $l = 0$, we have

$$\sigma^{(1)}(1) = \rho_0^{(1)} = \rho_0^{(2)} = \sigma^{(2)}(1).$$

LEMMA 2: *For $\mu > 1$ and $0 \leqq l < \mu - 1$, we have*

$$\sum_{1\leqq\mu_1<\cdots<\mu_l<\mu} q(\mu_1 \cdots \mu_l, \mu) < \frac{l+1}{\mu - l - 1}\sum_{1\leqq\mu_1<\cdots<\mu_{l+1}<\mu} q(\mu_1 \cdots \mu_{l+1}, \mu).$$

PROOF: We have, for any $\nu < \mu$, $\nu \neq \mu_i$ $(i = 1, \cdots, l)$

$$q(\mu_1 \cdots \mu_l\nu, \mu) > q(\mu_1 \cdots \mu_l, \mu).$$

Summing with respect to all such ν's,

$$\sum_\nu q(\mu_1 \cdots \mu_l\nu, \mu) > (\mu - l - 1)q(\mu_1 \cdots \mu_l, \nu).$$

Summing with respect to all $1 \leqq \mu_1 < \cdots < \mu_l < \mu$,

$$\sum_{1\leqq\mu_1<\cdots<\mu_l<\mu}\sum_\nu q(\mu_1 \cdots \mu_l\nu, \mu) = (l+1)\sum_{1\leqq\mu_1<\cdots<\mu_{l+1}<\mu} q(\mu_1 \cdots \mu_{l+1}, \mu)$$
$$> (\mu - l - 1)\sum_{1\leqq\mu_1<\cdots<\mu_l<\mu} q(\mu_1 \cdots \mu_l, \mu).$$

The lemma is proved.

Now we use induction to prove that for $\mu > 1$ and $l = 1, \cdots, \mu - 1$

$$\rho_l^{(1)} - \rho_l^{(2)} > \frac{d_0 + \binom{\mu-1}{1}d_1 + \cdots + \binom{\mu-1}{l}d_l}{\binom{\mu-1}{l}}$$
$$\times \sum_{1\leqq\mu_1<\cdots<\mu_l<\mu} q(\mu_1 \cdots \mu_l, \mu) \geqq 0.$$

This inequality holds for $l = 1$ by Lemma 2. Assume that it holds for l, $(l < \mu - 1)$. Then we have, by (5) and the fact that each $q < 0$,

$$\rho_{l+1}^{(1)} - \rho_{l+1}^{(2)} = \rho_l^{(1)} - \rho_l^{(2)} + d_{l+1} \sum_{1 \leqq \mu_1 < \cdots < \mu_{l+1} < \mu} q(\mu_1 \cdots \mu_{l+1}, \mu)$$

$$> \frac{d_0 + \binom{\mu-1}{1} d_1 + \cdots + \binom{\mu-1}{l} d_l}{\binom{\mu-1}{l}} \sum q(\mu_1 \cdots \mu_l, \mu) + d_{l+1} \sum q(\mu_1 \cdots \mu_{l+1}, \mu)$$

$$\geqq \left(\frac{d_0 + \binom{\mu-1}{1} d_1 + \cdots + \binom{\mu-1}{l} d_l}{\binom{\mu-1}{l}} \frac{l+1}{\mu - l - 1} + d_{l+1} \right) \sum q(\mu_1 \cdots \mu_{l+1}, \mu)$$

$$= \frac{d_0 + \binom{\mu-1}{1} d_1 + \cdots + \binom{\mu-1}{l} d_l + \binom{\mu-1}{l+1} d_{l+1}}{\binom{\mu-1}{l+1}} \sum q(\mu_1 \cdots \mu_{l+1}, \mu) \geqq 0.$$

Therefore, for $\mu > 1$, we have

$$\sigma^{(1)}(\mu) - \sigma^{(2)}(\mu) = \rho_{\mu-1}^{(1)} - \rho_{\mu-1}^{(2)} > 0.$$

Since $n > 1$ and $1 \leqq \mu \leqq n$, there exists a $\mu > 1$. Hence

$$\sum_{\mu=1}^{n} \sigma^{(1)}(\mu) > \sum_{\mu=1}^{n} \sigma^{(2)}(\mu)$$

which is equivalent to the inequality (1).

3. Our next step will be to obtain a generalization of Theorem 1. Consider a derived event defined by a disjunction of a (finite) number of events in the set, as follows:

$$E_{\nu_1} + E_{\nu_2} + \cdots + E_{\nu_m}.$$

We call such a disjunction a disjunction of the m-th order.

DEFINITION 3: *A set of events is said to be strongly mutually favorable in the second sense if for every positive integer m, the derived set of events consisting of all the disjunctions of the m-th order forms a strongly mutually favorable set of events (in the first sense).*

Let $D = D(m)$ denote in general a disjunction of the m-th order; let $p(D_1 \cdots D_h, D)$ denote the probability of the occurrence of the disjunction D, on the hypothesis that the conjunction of the h disjunctions $D_1 \cdots D_h$ has occurred. Then Definition 3 says that for any positive integer h and any set of distinct D's we have

$$p(D_1 \cdots D_h, D) > p(D_1 \cdots D_{h-1}, D).$$

Since a disjunction of the 1st order is an event E, we see that Definition 3 includes Definition 2.

Let $D_m(\nu_1, \cdots, \nu_k)$, $\nu_1 < \cdots < \nu_k$ denote the derived event

$$\prod_{\mu_1, \cdots, \mu_m} (E_{\mu_1} + \cdots + E_{\mu_m})$$

where the product (conjunction) extends to all combinations of m indices out of the indices $\nu_1, \cdots, \nu_k$. Let $p_m^*(\nu_1, \cdots, \nu_k)$ denote the probability of the occurrence of $D_m(\nu_1, \cdots, \nu_k)$. It is seen that $p_1^*(\nu_1, \cdots, \nu_k) = p(\nu_1 \cdots \nu_k)$ in our previous notation.

We merely state Theorem 2, whose proof is analogous to that of Theorem 1 but requires more cumbersome expressions.

THEOREM 2: *Let $n > k \geqq m \geqq 1$ and let $E_1, \cdots, E_n$ be a set of mutually strongly favorable events in the second sense. Then we have*

$$\prod_{1 \leqq \nu_1 < \cdots < \nu_{k+1} \leqq n} [p_m^*(\nu_1, \cdots, \nu_{k+1})]^{\binom{n-m}{k-m+1}^{-1}} > \prod_{1 \leqq \nu_1 < \cdots < \nu_k \leqq n} [p_m^*(\nu_1, \cdots, \nu_k)]^{\binom{n-m}{k-m}^{-1}}.$$

To give an interpretation of $p_m^*(\nu_1, \cdots, \nu_k)$, we prove the symbolic equation between events:

$$D_m = \prod_{\nu_1 \leqq \mu_1 < \cdots < \mu_m \leqq \nu_k} (E_{\mu_1} + \cdots + E_{\mu_m}) = \sum_{\nu_1 \leqq \mu_1 < \cdots < \mu_{k-m+1} \leqq \nu_k} (E_{\mu_1} \cdots E_{\mu_{k-m+1}}) = C_{k-m+1},$$

where product means conjunction and sum means disjunction.

To prove this, we write for a general event E, $E = 1$ when E occurs, $E = 0$ when E does not occur. Now if $C_{k-m+1} = 0$, then at most $k - m$ events among the k given events occur, so that there exist m events such that $E_{\lambda_1} = 0$, $E_{\lambda_2} = 0$, $E_{\lambda_m} = 0$, thus

$$E_{\lambda_1} + E_{\lambda_2} + \cdots + E_{\lambda_m} = 0.$$

Now the last disjunction is contained in D_m as a factor, therefore $D_m = 0$.

Conversely, if $D_m = 0$, at least one of its factors $= 0$, so that there exist m events, such that $E_{\lambda_1} = 0$, $E_{\lambda_2} = 0$, $\cdots$, $E_{\lambda_m} = 0$. Thus at most $k - m$ events out of the k given events occur and so by definition $C_{k-m+1} = 0$. Q.e.d.

From the above it immediately follows that

$$p_m^*(\nu_1, \cdots, \nu_k) = p_{k-m+1}(\nu_1, \cdots, \nu_k)$$

where $p_{k-m+1}(\nu_1, \cdots, \nu_k)$ is defined in the Introduction. Then Theorem 2 may be written as

$$\Pi[p_{k-m+2}(\nu_1, \cdots, \nu_{k+1})]^{\binom{n-m}{k-m+1}^{-1}} > \Pi[p_{k-m+1}(\nu_1, \cdots, \nu_k)]^{\binom{n-m}{k-m}^{-1}}$$

or again as

$$\Pi[w_{m-1}(\nu_1', \cdots, \nu_{k+1}')]^{\binom{n-m}{k-m+1}^{-1}} > \Pi[w_{m-1}(\nu_1', \cdots, \nu_k')]^{\binom{n-m}{k-m}^{-1}}$$

where $w_{m-1}(\nu_1', \cdots, \nu_k')$ denotes the probability of the occurrence of at most $m - 1$ events out of the k events $E'_{\nu_1}, \cdots, E'_{\nu_k}$.

REMARK. If in our Definitions 2 and 3 we replace the sign "$>$" by the sign "$\geqq$", then we obtain the inequalities in Theorems 1 and 2 with the sign "$>$" replaced by "$\geqq$". The corresponding set of events thus newly defined will be said to be strongly mutually favorable or indifferent (in the first or second sense).

After this modification, we can include events with the probability 1 in our considerations. Also, the events need no longer be distinct and there may now exist implication relations between events or their conjunctions. This modification is useful for the following application.

4. Consider the divisibility of a random positive integer by the set of positive integers. To each positive integer there corresponds an event, namely the event that the random positive integer is divisible by it. The enumerable set of events

$$E_1, E_2, E_3, E_4, \cdots, E_n, \cdots$$

where E_n = the event of divisibility by n, with the probabilities

$$1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \cdots, \frac{1}{n}, \cdots$$

evidently forms a set of strongly mutually favorable or indifferent events in the second sense.

Again, the enumerable set of events

$$E'_2, E'_3, E'_4, \cdots, E'_n, \cdots$$

where E'_n = the event of non-divisibility by n, with the probabilities

$$\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \cdots, \frac{n-1}{n}, \cdots$$

evidently also forms a set of strongly mutually favorable or indifferent events in the second sense.

Hence our Theorem 2 can be applied to both sets and in this way we obtain results which belong properly to the elementary theory of numbers.

We shall content ourselves with indicating a few examples.

Let $\{a_1, \cdots, a_n\}$ denote the least common multiple of the natural numbers $a_1, \cdots, a_n$. Then Theorem 1, when applied to the two sets above, gives respectively

THEOREM 1.1: *Let $a_1, \cdots, a_n$ be any positive integers, then we have, for $k = 1, \cdots, n - 1$*

$$\left(\prod_{1 \leqq \nu_1 < \cdots < \nu_{k+1} \leqq n} \frac{1}{\{a_{\nu_1}, \cdots, a_{\nu_{k+1}}\}}\right)^{\binom{n-1}{k}^{-1}} \geqq \left(\prod_{1 \leqq \nu_1 < \cdots < \nu_k \leqq n} \frac{1}{\{a_{\nu_1}, \cdots, a_{\nu_k}\}}\right)^{\binom{n-1}{k-1}^{-1}}.$$

THEOREM 1.2: *Also we have,*

$$\prod_{1\le \nu_1<\cdots<\nu_{k+1}\le n}\left(1-\sum_{\nu_1\le\mu_1\le\nu_{k+1}}\frac{1}{a_{\mu_1}}+\sum_{\nu_1\le\mu_1<\mu_2\le\nu_{k+1}}\frac{1}{\{a_{\mu_1},a_{\mu_2}\}}-+\cdots+(-1)^{k+1}\frac{1}{\{a_{\nu_1},\cdots,a_{\nu_{k+1}}\}}\right)^{\binom{n-1}{k}^{-1}}$$
$$\ge\prod_{1\le \nu_1<\cdots<\nu_{k}\le n}\left(1-\sum_{\nu_1\le\mu_1\le\nu_{k}}\frac{1}{a_{\mu_1}}+\sum_{\nu_1\le\mu_1<\mu_2\le\nu_{k}}\frac{1}{\{a_{\mu_1},a_{\mu_2}\}}-+\cdots+(-1)^{k}\frac{1}{\{a_{\nu_1},\cdots,a_{\nu_{k}}\}}\right)^{\binom{n-1}{k-1}^{-1}}$$

A trivial corollary of Theorem 1 is

$$p(12\cdots n)\ge p_1p_2\cdots p_n\,.$$

Correspondingly we have

$$1-\sum_{1\le\mu_1\le n}\frac{1}{a_{\mu_1}}+\sum_{1\le\mu_1<\mu_2\le n}\frac{1}{\{a_{\mu_1},a_{\mu_2}\}}-+\cdots+(-1)^n\frac{1}{\{a_1,\cdots,a_n\}}$$
$$\ge\left(1-\frac{1}{a_1}\right)\left(1-\frac{1}{a_2}\right)\cdots\left(1-\frac{1}{a_n}\right).$$

If we multiply by $a_1a_2\cdots a_n$, we get

$$A(a_1,a_2,\cdots,a_n)\ge(a_1-1)(a_2-1)\cdots(a_n-1),$$

where $A(a_1,\cdots,a_n)$ denotes the number of positive integers $\le a_1a_2\cdots a_n$ that are not divisible by any of the a_i $(i=1,\cdots,n)$.

This last result, which is almost obvious here, was proved by H. Rohrbach and H. Heilbronn independently.[2] See also my generalization[3] (also obvious from the present point of view) of this result to higher dimensional sets of positive integers and to sets of ideals in any algebraic number field.

[2] "Beweis einer zahlentheoretische Ungleichung," *Jour. für Math.*, Vol. 177 (1937), pp. 193-196. "On an inequality in the elementary theory of numbers," *Proc. Camb. Phil. Soc.*, Vol. 33, (1937), pp. 207-209.

[3] "A generalization of an inequality in the elementary theory of numbers," *Jour. für Math.*, Vol. 183 (1941), p. 103.

ANNALS OF MATHEMATICS
Vol. 48, No. 4, October, 1947

ON THE LOWER LIMIT OF SUMS OF INDEPENDENT RANDOM VARIABLES

BY KAI-LAI CHUNG AND PAUL ERDÖS

(Received October 30, 1946)

1. Let $X_1, X_2, \cdots, X_n \cdots$ be independent random variables and let $S_n = \sum_{\nu=1}^{n} X_\nu$. In the so-called law of the iterated logarithm, completely solved by Feller recently, the upper limit of S_n as $n \to \infty$ is considered and its true order of magnitude is found with probability one. A counterpart to that problem is to consider the lower limit of S_n as $n \to \infty$ and to make a statement about its order of magnitude with probability one.

THEOREM 1. *Let $X_1, \cdots, X_n, \cdots$ be independent random variables with the common distribution*: $\Pr(X_n = 1) = p, \Pr(X_n = 0) = 1 - p = q$. *Let* $\psi(n) \downarrow \infty$ and

$$\sum_{n=1}^{\infty} \frac{1}{n\psi(n)} = \infty. \tag{1.1}$$

Then we have

$$\Pr\left(\varliminf_{n\to\infty} n^{1/2}\psi(n)\,|S_n - np| = 0\right) = 1. \tag{1.2}$$

Theorem 1 is a best possible theorem. In fact we shall prove the following

THEOREM 2. *Let X_n be as in Theorem 1 but let p be a quadratic irrational. Let $\phi(n) \uparrow \infty$ and*

$$\sum_{n=1}^{\infty} \frac{1}{n\phi(n)} < \infty. \tag{1.3}$$

Then we have

$$\Pr\left(\varliminf_{n\to\infty} n^{1/2}\phi(n)\,|S_n - np| = 0\right) = 0. \tag{1.4}$$

By making use of results on uniform distribution mod 1 we can prove (1.4) for almost all p, however the proof is omitted here.

In order to extend the theorem to more general sequences of random variables, we need a theorem about the limiting distribution of S_n with an estimate of the accuracy of approximation. Cramér's asymptotic expansion is suitable for this purpose. The conditions on $F(x)$ in the following Theorem 3 are those under which the desired expansion holds.

THEOREM 3. *Let $X_1, \cdots, X_n, \cdots$ be independent random variables having the same distribution function $F(x)$. Suppose that the absolutelv continuous part of $F(x)$ does not vanish identically and that its first moment is zero, the second is one, and the absolute fifth is finite. Let $\psi(n)$ be as in Theorem 1, then*

$$\Pr\left(\varliminf_{n\to\infty} n^{1/2}\psi(n)\,|S_n| = 0\right) = 1. \tag{1.5}$$

On the other hand, let $\phi(n)$ be as in Theorem 2; then

$$\Pr\left(\varliminf_{n\to\infty} n^{1/2}\,\phi(n)\,|\,S_n\,| = 0\right) = 0. \tag{1.6}$$

It seems clear that the result can be extended to other cases, however we shall at present content ourselves with this statement.

2. For $x > 0$ let $I(x)$ denote the integer nearest to x if x is not equal to $[x] + \frac{1}{2}$; in the latter case, let $I(x) = [x]$; let $\{x\} = x - I(x)$. We have then for any $x > 0$, $y > 0$, the inequality

$$|\,\{x - y\}\,| \leqq |\,\{x\} - \{y\}\,|.$$

We are now going to state and prove some lemmas. The first two lemmas are number-theoretic in nature; the third one supplies the main probability argument; and the fourth one is a form of zero-or-one law.

LEMMA 1. *Let $p > 0$ be a real number. Let $\psi(n) \uparrow \infty$. Arrange all the positive integers n for which we have,*

$$|\,\{np\}\,| < cn^{-1/2}\psi(n)^{-1} \tag{2.1}$$

in an increasing sequence n_i, $i = 1, 2, \cdots$. Then for any pair of positive integers i and k we have

$$n_{i+2k} \geqq n_i + n_k.$$

PROOF. Suppose the contrary:

$$n_{i+2k} < n_i + n_k.$$

CASE (i): $k \leqq i$. Consider the $2k + 1$ numbers

$$n_i,\ n_{i+1},\ \cdots,\ n_{i+2k}$$

and the corresponding

$$\{n_ip\},\ \{n_{i+1}p\},\ \cdots,\ \{n_{i+2k}p\}. \tag{2.2}$$

There are at least $k + 1$ numbers among (2.2) which are of the same sign; without loss of generality we may assume that they are non-negative. Let the corresponding n_i be

$$n_{i_1} < n_{i_2} < \cdots < n_{i_{k+1}}.$$

Then we have

$$0 \leqq \{n_{i_j}\,p\} < cn_{i_j}^{-1/2}\psi(n_i)^{-1} \leqq cn_k^{-1/2}\psi(n_k)^{-1}, \qquad j = 1, \cdots, k+1;$$

since $i_j \geqq i \geqq k$; and

$$|\,\{n_{i_{k+1}}\,p - n_{i_j}\,p\}\,| < cn_k^{-1/2}\psi(n_k)^{-1}, \qquad j = 1, \cdots, k;$$

$$0 < n_{i_{k+1}} - n_{i_j} \leqq n_{i+2k} - n_i < n_k.$$

Thus there would be k different positive integers $n_{i_{k+1}} - n_{i_j}$, $j = 1, \cdots, k$ all $< n_k$, for which

$$| \{np\} | < cn_k^{-1/2}\psi(n_k)^{-1}.$$

This is a contradiction to the definition of n_k.

CASE (ii) $k > i$. Consider the $i + k + 1$ numbers

$$n_k, \ n_{k+1}, \ \cdots, \ n_{i+2k}$$

and the corresponding

$$\{n_k p\}, \{n_{k+1}p\}, \cdots, \{n_{i+2k}p\}.$$

Since $i + k + 1 > 2i + 1$, there are at least $i + 1$ of the numbers above which are of the same sign, say non-negative. Let the corresponding n_i be

$$n_{k_1} < n_{k_2} < \cdots < n_{k_{i+1}}.$$

By an argument similar to that in Case (i) we should have i numbers $n_{k_{i+1}} - n_{k_j}$, $j = 1, \cdots, i$, all $< n_i$ for which

$$| \{np\} | < cn_i^{-1/2}\psi(n_i)^{-1}.$$

This leads to a contradiction as before.

LEMMA 2. *Let n_i be defined as in Lemma 1. Then if*

$$\sum_{n=1}^{\infty} \frac{1}{n\psi(n)} = \infty, \tag{2.3}$$

we have

$$\sum_{i=1}^{\infty} n_i^{-1/2} = \infty. \tag{2.4}$$

PROOF. Consider the points

$$hcn^{-1/2}\psi(n)^{-1} \qquad\qquad h = \pm 1, \cdots, \pm[2^{-1}c^{-1}n^{1/2}\psi(n)].$$

They divide the interval $(-\frac{1}{2}, \frac{1}{2})$ into at most $[c^{-1}n^{1/2}\psi(n)] + 2$ parts. Hence at least one subinterval contains

$$l \geq \frac{n}{[c^{-1}n^{1/2}\psi(n)] + 2}$$

members of the n numbers $\{mp\}$, $m = 1, 2, \cdots, n$. Let the corresponding n_i be

$$n_1 < n_2 < \cdots < n_l.$$

Then

$$0 < | \{n_l p - n_i p\} | < cn^{-1/2}\psi(n)^{-1} < c(n_l - n_i)^{-1/2}\psi(n_l - n_i)^{-1}, \qquad i = 1, \cdots, l - 1.$$

Hence if $g(n)$ denote the number of numbers among $1, \cdots, n$ for which

$$| \{np\} | < cn^{-1/2}\psi(n)^{-1},$$

we have, for n sufficiently large

$$g(n) > 2^{-1}cn^{1/2}\psi(n)^{-1}.$$

Now

$$\sum_{2^{k-1}<n_i\leq 2^k} n_i^{-1/2} \geq \frac{g(2^k) - g(2^{k-1})}{\sqrt{2^k}}.$$

Hence

$$\begin{aligned}\sum_{k=1}^{\infty} \sum_{2^{k-1}<n_i\leq 2^k} n_i^{-1/2} &\geq \sum_{k=1}^{\infty} \frac{g(2^k) - g(2^{k-1})}{\sqrt{2^k}} \\ &= -\frac{g(1)}{\sqrt{2}} + \sum_{k=1}^{\infty} g(2^k)\left(\frac{1}{\sqrt{2^k}} - \frac{1}{\sqrt{2^{k+1}}}\right) \\ &\geq -\frac{g(1)}{\sqrt{2}} + \sum_{k=1}^{\infty} \frac{c}{2}\frac{\sqrt{2^k}}{\psi(2^k)}\left(\frac{1}{\sqrt{2^k}} - \frac{1}{\sqrt{2^{k+1}}}\right) \\ &\geq -1 + \frac{c}{2}\left(1 - \frac{1}{\sqrt{2}}\right)\sum_{k=1}^{\infty}\frac{1}{\psi(2^k)}.\end{aligned}$$

It is well-known[1] that if (2.3) holds then

$$\sum_{k=1}^{\infty} \frac{1}{\psi(2^k)} = \infty.$$

Thus (2.4) is proved.

LEMMA 3. *Let n_i, $i = 1, 2, \cdots$ be a monotone increasing sequence such that for any pair of positive integers i and k we have*

$$n_{i+2k} \geq n_i + n_k \tag{2.5}$$

and

$$\sum_{i=1}^{\infty} n_i^{-1/2} = \infty. \tag{2.6}$$

Then if α and β are two integers, we have for any integer $h > 0$,

$$\Pr(S_{n_i} = I(pn_i + p\alpha) + \beta \text{ at least once for } i \geq h) \geq \tfrac{1}{2}. \tag{2.7}$$

[1] See e. g. Theory and Application of Infinite Series, London-Glasgow, Blackie and Son, 1928, p. 120.

PROOF. Denoting the joint probability of E_1, E_2, $\cdots$ by $\Pr(E_1; E_2; \cdots)$ we have

$$\Pr(S_{n_i} = I(pn_i + p\alpha) + \beta) = \Pr(S_{n_h} = I(pn_h + p\alpha) + \beta;$$
$$S_{n_i - n_h} = I(pn_i + p\alpha) - I(pn_h + p\alpha))$$
$$+ \Pr(S_{n_h} \neq I(pn_h + p\alpha) + \beta; S_{n_{h+1}} = I(pn_{h+1} + p\alpha) + \beta;$$
$$S_{n_i - n_{h+1}} = I(pn_i + p\alpha) - I(pn_{h+1} + p\alpha))$$
$$+ \cdots$$
$$+ \Pr(S_{n_h} \neq I(pn_h + p\alpha) + \beta; \cdots; S_{n_{i-1}} \neq I(pn_{i-1} + p\alpha) + \beta;$$
$$S_{n_i} = I(pn_i + p\alpha) + \beta).$$

Writing

$$p_i = \Pr(S_{n_i} = I(pn_i + p\alpha) + \beta),$$
$$w_k = \Pr(S_{n_j} \neq I(pn_j + p\alpha) + \beta \quad \text{for} \quad h \leqq j < k; S_{n_k} = I(pn_k + p\alpha) + \beta),$$
$$p_{k,i} = \Pr(S_{n_i - n_k} = I(pn_i + p\alpha) - I(pn_k + p\alpha)), \qquad p_{k,k} = 1;$$

and using the assumption of independence, we have

$$p_i = \sum_{k=h}^{i} w_k p_{k,i}.$$

Summing from h to m we get

$$\sum_{i=h}^{m} p_i = \sum_{i=h}^{m} \sum_{k=1}^{i} w_k p_{k,i} \leqq \sum_{k=1}^{m} w_k \sum_{i=k}^{m} p_{k,i}. \tag{2.8}$$

Now for any positive x and y, $I(x) - I(y) = I(x - y)$ or $I(x - y) \pm 1$; and it is well-known that for the random variables we have, given any $\epsilon > 0$, if $n > n_0(\epsilon)$, and $\theta = \pm 1$,

$$\Pr(S_n = I(np) + \theta) \leqq (1 + \epsilon)\Pr(S_n = I(np))$$

hence we have, if $i - k \geqq m_1(\epsilon)$,

$$p_{k,i} \leqq (1 + \epsilon/4)\Pr(S_{n_i - n_k} = I(pn_i - pn_k)). \tag{2.9}$$

From (2.5) if $i > k$, we have

$$n_i \geqq n_k + n_{[(i-k)/2]}. \tag{2.10}$$

Also it is well-known that as $i \to \infty$,

$$p_i \sim \frac{1}{\sqrt{2\pi pq n_i}}. \tag{2.11}$$

Hence from (2.9), (2.10) and (2.11) we have if $i - k \geqq m_2(\epsilon)$ where m_2 is a positive constant,

$$p_{k,i} \leqq (1 + \epsilon/2)\Pr(S_{n_{[(i-k)/2]}} = I(pn_{[(i-k)/2]})).$$

Since α and β are fixed, to any $\epsilon > 0$ there exists an integer $m_0 = m_0(\epsilon) > m_2$ such that if $i - k \geqq m_0(\epsilon)$,

$$\Pr(S_{n_{[(i-k)/2]}} = I(pn_{[(i-k)/2]})) \leqq (1 + \epsilon)p_{[(i-k)/2]}. \tag{2.12}$$

Thus for $i - k \geqq m_0(\epsilon)$,

$$p_{k,i} \leqq (1 + \epsilon)p_{[(i-k)/2]}.$$

Using (2.12) in (2.9), we obtain

$$\sum_{i=h}^{\infty} p_i \leqq \sum_{k=1}^{\infty} w_j \left(\sum_{i=k}^{k+m_0-1} p_{j,i} + (1 + \epsilon) \sum_{i=k+m_0}^{m} p_{[(i-k)/2]} \right)$$
$$\leqq \sum_{k=1}^{m} w_j \left(m_0 + 2(1 + \epsilon) \sum_{i=m_0}^{[m/2]} p_i \right).$$

Therefore

$$\sum_{j=1}^{m} w_j \geqq \frac{\sum_{i=h}^{m} p_i}{m_0 + 2(1 + \epsilon) \sum_{i=m_0}^{[m/2]} p_i}.$$

Since by (2.11) and (2.6) the series $\sum_{i=1}^{\infty} p_i$ is divergent, we get, letting $n \to \infty$,

$$\sum_{j=1}^{\infty} w_j \geqq \frac{1}{2(1 + \epsilon)}.$$

Since ϵ is arbitrary and the left-hand side does not depend on ϵ this proves (2.7).

LEMMA 4. *If for any integers α, β and $k > 0$, there exists a number $\eta > 0$ not depending on α, β and an integer $l = l(k, \eta)$ such that, n_i being any sequence $\uparrow \infty$,*

$$\Pr(S_{n_i} = I(pn_i + p\alpha) + \beta \text{ at least once for } k \leqq i \leqq l) \geqq \eta; \tag{2.13}$$

then

$$\Pr(S_{n_i} = I(pn_i + p\alpha) + \beta \text{ infinitely often}) = 1. \tag{2.14}$$

PROOF. Take a sequence $k_1, k_2, \cdots$ and the corresponding $l_1, l_2, \cdots$ such that

$$k_1 < l_1 < k_2 < l_2 < \cdots$$

Consider the event

$$E_r: \qquad S_{n_i} = I(pn_i + p\alpha) + \beta \text{ at least once for } k_r \leqq i \leqq l_r,$$

and let the probability that E_r occurs under the hypothesis that none of $E_1, \cdots, E_{r-1}$ occurs, be denoted by $\Pr(E_r \mid E_1' \cdots E_{r-1}')$. Then the latter is a probability mean of the conditional probabilities of E_r under the various hypotheses:

$$\text{H:} \qquad S_{n_i} = \sigma_{n_i}, \quad k_t \leqq i \leqq l_t, \quad 1 \leqq t \leqq r - 1;$$

where the σ_{n_i}'s are such that for all i, $\sigma_{n_i} \neq I(pn_i + p\alpha) + \beta$ but are otherwise arbitrary. Now under H, E_r will occur if the following event F occurs:

F: $S_{n_i - n_{l_{r-1}}} = I(pn_i + p\alpha) + \beta - \sigma_{n_{l_{r-1}}}$ at least once for $k_r \leq i \leq l_r$.

Hence

$$\Pr(E_r \mid E_1' \cdots E_{r-1}') \geqq \min_H \Pr(E_r \mid H) \geqq \Pr(F \mid H) = \Pr(F).$$

Writing the equality in F as

$$\begin{aligned} S_{n_i - n_{l_{r-1}}} &= I(p(n_i - n_{l_{r-1}}) + p(n_{l_{r-1}} + \alpha)) + \beta - \sigma_{n_{l_{r-1}}} \\ &= I(p(n_i - n_{l_{r-1}}) + p\alpha') + \beta' \end{aligned}$$

and consider the random variables $X_{n_{l_{r+1}}}, X_{n_{l_{r+1}}+1}, \cdots$ as $X_1', X_2', \cdots$ we see from (2.13) that

$$\Pr(E_r \mid E_1' \cdots E_{r-1}') \geqq \Pr(F) \geqq \eta.$$

Therefore the probability that none of the events E_r, $r = 1, \cdots, s$ occurs is $\Pr(E_1' \cdots E_s') = \Pr(E_1')\Pr(E_2' \mid E_1') \cdots \Pr(E_s' \mid E_1' \cdots E_{s-1}') \leqq (1 - \eta)^s$. Hence

$$\Pr(S_{n_i} \neq I(pn_i + p\alpha) + \beta \text{ for all } l_r \leqq i \leqq k_r, r = 1, 2, \cdots) = 0$$

Since l_1 can be taken arbitrarily large, (2.14) is proved.

REMARK. Lemma 3 and 4 imply an interesting improvement of the well-known fact that $\Pr(S_n - np = \text{infinitely often}) = 1$ for a rational p. Let n_i be any monotone increasing sequence such that (2.6) holds; in addition if for a certain integer $m > 0$ and any pair of integers i and k we have

$$n_{i+mk} \geqq n_i + n_k \tag{2.15}$$

then

$$\Pr(S_{n_i} - n_i p = 0 \text{ for infinitely many } i) = 1.$$

That the condition (2.6) alone is not sufficient can be shown by a counter-example. On the other hand, it is trivial that (2.6) is a necessary condition. The condition (2.15) can be replaced e.g. by the following condition:

$$n_{i+1} - n_i \geqq A n_i^{1/2}, \qquad A > 0.$$

The proof is different and will be omitted here.

PROOF OF THEOREM 1. Let the sequence n_i be defined as in Lemma 1. Then by Lemma 1 and 2 this sequence satisfies the conditions (2.5) and (2.6) in Lemma 3. Hence by Lemma 3 the condition (2.13) in Lemma 4 is satisfied with any $\eta < \frac{1}{2}$. Thus byLemma 4 we have (2.14). Taking $\alpha = \beta = 0$ therein we obtain

$$\Pr(S_{n_i} - n_i p = \{n_i p\} \text{ infinitely often}) = 1.$$

Hence by the definition (2.2)

$$\Pr(|S_n - np| < cn^{-1/2}\psi(n)^{-1} \text{ infinitely often}) = 1.$$

Since c is arbitrarily small (1.2) is proved.

REMARK. It is clear that (2.14) yields more than Theorem 1 since α and β are arbitrary. It is easily seen that we may even make α and β vary with n_i in a certain way, but we shall omit these considerations here.

PROOF OF THEOREM 2. Arrange all the positive integers n for which we have

$$|\{np\}| \leq An^{-1/2}\phi(n)^{-1}, \qquad A > 0.$$

in an ascending sequence n_i, $i = 1, 2, \cdots$. Since

$$|\{n_i p\}| \leq An_i^{-1/2}\phi(n_i)^{-1}$$

we have

$$|\{n_{i+1}p - n_i p\}| \leq 2An_i^{-1/2}\phi(n_i)^{-1}. \tag{2.16}$$

On the other hand, since p is a quadratic irrational, it is well-known[2] that there exists a number $M > 0$ such that

$$|\{n_{i+1}p - n_i p\}| > \frac{M}{n_{i+1} - n_i}. \tag{2.17}$$

From (2.16) and (2.17) we get with $A_1 = M/2A$,

$$n_{i+1} - n_i > A_1 n_i^{1/2}\phi(n_i) \tag{2.18}$$

Without loss of generality we may assume that $\phi(n_i) \leq n_i^{1/2}$. For we may replace $\phi(n)$ by $\phi_1(n)$ defined as follows:

$$\phi_1(n) = \begin{cases} \phi(n) & \text{if} \quad \phi(n) \leq n^{1/2}; \\ n^{1/2} & \text{if} \quad \phi(n) > n^{1/2}. \end{cases}$$

After this replacement (1.3) remains convergent, while if (1.4) holds for $\phi_1(n)$, it holds *a fortiori* for $\phi(n)$.

Now if $\phi(n_i) \leq n_i^{1/2}$, and the constant A_2 is such that $2A_2 + A_2^2 < A_1$, we have from (2.18)

$$n_{i+1}^{1/2} > n_i^{1/2} + A_2\phi(n_i).$$

Hence by iterating,

$$n_{i+1}^{1/2} > A_2 \sum_{k=1}^{i} \phi(n_k) > A_2 \sum_{k=[i/2]}^{i} \phi(n_k) > A_2 \frac{i}{2}\phi\left(\left[\frac{i}{2}\right]\right).$$

Therefore by (1.3)

$$\sum_{i=1}^{\infty} n_i^{-1/2} < \infty. \tag{2.19}$$

[2] See e. g. HARDY AND WRIGHT, Introduction to the Theory of Numbers, Oxford 1938, p. 157.

Define

$$p_i = \Pr(S_{n_i} = I(pn_i)).$$

As in (2.11) we have

$$p_i \sim \frac{1}{\sqrt{2\pi pqn_i}}.$$

Hence from (2.18)

$$\sum_{i=1}^{\infty} p_i < \infty.$$

By the classical Borel-Cantelli lemma it follows that

$$\Pr(S_{n_i} = I(pn_i) \text{ infinitely often}) = 0.$$

By the definition of n_i this is equivalent to (1.4).

3. LEMMA 5. *Let $X_1, \cdots, X_n, \cdots$ be independent random variables having the same distribution function $F(x)$ which satisfies the conditions in Theorem 3. Then if $x_1 < x_2$ and $x_1 = o(1)$, $x_2 = o(1)$ as $n \to \infty$, we have*

(3.1) $$\Pr(x_1 \leqq n^{-1/2} S_n \leqq x_2) = (2\pi)^{-1/2}(x_2 - x_1) + o(x_2 - x_1) + 0(n^{-3/2})$$

PROOF. By Cramér's asymptotic expansion[3] we have, if we denote the r[th] moment of $F(x)$ by α_r,

$$\Pr\left(\frac{S_n}{\sqrt{n}} \leqq x\right) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-y^2/2}\,dy - \frac{\alpha_3}{6\sqrt{2\pi}\sqrt{n}}(x^2 - 1)\,e^{-x^2/2}$$
$$+ \frac{\alpha_4 - 3\alpha_2^2}{24\sqrt{2\pi}\,n}(-x^3 + 3x)e^{-x^2/2} + \frac{\alpha_3^2}{72\sqrt{2\pi}\,n}(-x^5 + 10x^3 - 15x)e^{-x^2/2} + R(x)$$

where

$$|R(x)| \leqq Qn^{-3/2},$$

and Q is a constant depending only on $F(x)$,

It follows, using elementary estimates, that

$$\Pr\left(x_1 \leqq \frac{S_n}{\sqrt{n}} \leqq x_2\right) = \frac{1}{\sqrt{2\pi}}\int_{x_1}^{x_2} e^{-y^2/2}\,dy$$
$$+ 0\left((x_2 - x_1)\left(\frac{|x_1| + |x_2|}{\sqrt{n}} + \frac{1}{n}\right)\right) + 0\left(\frac{1}{\sqrt{n^3}}\right)$$

Since $x_1 = o(1)$, $x_2 = o(1)$ this reduces immediately to (3.1).

[3] CRAMÉR, Random Variables and Probability Distributions, Cambridge 1937, Ch. 7. For a simplified proof see P. L. HSU, *The Approximate Distribution of the Mean and Variance of a Sample of Independent Variables*, Ann. Math. Statistics, 16 (1945), pp. 1-29.

LEMMA 6. *Let z_n be any real number such that $z_n = O(n^{1/2})$, c any positive number, and h any positive integer. Let $\psi(n) \uparrow \infty$ and*

$$\sum_{n=1}^{\infty} \frac{1}{n\psi(n)} = \infty. \tag{3.2}$$

Then if the random variables X_n satisfy the conditions of Theorem 3, we have

$$\Pr(\,|S_n - z_n| \leqq cn^{-1/2}\psi(n)^{-1} \text{ at least once for } n \geqq h) = 1. \tag{3.3}$$

PROOF. Write

$$P_n = \Pr(|S_n - z_n| \leqq cn^{-1/2}\psi(n)^{-1});$$

$$W_k = \Pr(|S_j - z_j| > cj^{-1/2}\psi(j)^{-1} \text{ for } h \leqq j < k; \quad |S_k - z_k| \leqq ck^{-1/2}\psi(k)^{-1})$$

$$P_{k,n} = \Pr(|S_n - z_n| \leqq cn^{-1/2}\psi(n)^{-1} \mid |S_j - z_j| > cj^{-1/2}\psi(j)^{-1} \text{ for } h \leqq j < k; \; |S_k - z_k| \leqq ck^{-1/2}\psi(k)^{-1}).$$

Then by a similar argument as in Lemma 3, we have

$$\sum_{n=h}^{m} P_n \leqq \sum_{k=h}^{m} W_k \sum_{n=k}^{m} P_{k,n}. \tag{3.4}$$

Our next step is to show that to any $\epsilon > 0$ there exists a constant $A(\epsilon)$ such that for $n - k > A$, we have

$$P_{k,n} \leqq (1 + \epsilon)P_{n-k}. \tag{3.5}$$

To prove this we divide the x-interval $|x - z_k| \leqq ck^{-1/2}\psi(k)^{-1}$ into disjoint subintervals I_j; of lengths $\leqq \epsilon' cn^{-1/2}\psi(n)^{-1}$ where $\epsilon' > 0$ is arbitrary. If we write

$$P_{k,n}^{(j)} = \Pr(\,|S_n - z_n| \leqq cn^{-1/2}\psi(n)^{-1} \mid S_k - z_k \subset I_j)$$

we have

$$P_{k,n}^{(j)} \leqq \Pr(S_n - S_k \subset I_j')$$

where I_j' is an interval of lengths $\leqq (2 + \epsilon')cn^{-1/2}\psi(n)^{-1} \leqq (2 + \epsilon')c(n - k)^{-1/2}$ $\psi(n - k)^{-1}$ lying within the interval $|x - z_n + z_k| \leqq cn^{-1/2}\psi(n)^{-1} + ck^{-1/2}\psi(k)^{-1}$. From Lemma 5 it is seen that if $n - k \geqq A_1(\epsilon')$,

$$P_{k,n}^{(j)} \leqq \frac{2(1 + \epsilon')c}{\sqrt{2\pi}\,(n - k)\psi(n - k)};$$

since $P_{k,n}$ is a probability mean of $P_{k,n}^{(j)}$, we have

$$P_{k,n} \leqq \max_j P_{k,n}^{(j)} \leqq \frac{2(1 + \epsilon')c}{\sqrt{2\pi}\,(n - k)\psi(n - k)}. \tag{3.6}$$

On the other hand, we have again from Lemma 5, if $n - k \geqq A_2(\epsilon')$,

$$P_{n-k} \geqq \frac{2(1 - \epsilon')}{\sqrt{2\pi}\,(n - k)\psi(n - k)}. \tag{3.7}$$

From (3.6) and (3.7) follows (3.5).

Using (3.5) in (3.4) we get

$$\sum_{n=h}^{\infty} P_n \leqq \sum_{k=h}^{\infty} W_k \left(\sum_{n=k}^{k+A-1} P_{k,n} + (1+\epsilon) \sum_{n=k+A}^{m} P_{n-k} \right)$$
$$\leqq \sum_{k=h}^{m} W_k \left(A + (1+\epsilon) \sum_{n=A}^{m} P_n\right)$$

$$\sum_{k=h}^{m} W_k \geqq \frac{\sum_{n=h}^{m} P_n}{A + (1+\epsilon) \sum_{n=A}^{m} P_n} \tag{3.8}$$

Now $\sum_{n=h}^{\infty} P_n = \infty$ by (3.7) and (3.1). It follows from (3.8) by letting $n \to \infty$ that

$$\sum_{k=h}^{\infty} W_k \geqq \frac{1}{1+\epsilon}.$$

Since ϵ is arbitrary and the left-hand side does not depend on ϵ we have

$$\sum_{k=h}^{\infty} W_k \geqq 1. \tag{3.9}$$

Thus (3.3) follows.

PROOF OF THEOREM 3. Taking $z = 0$ in (3.9) and denoting by E_k the event

$$| S_n | \leqq cn^{-1/2}\psi(n)^{-1},$$

we can write (3.9) as follows:

$$\Pr\left(\sum_{n=h}^{\infty} E_n\right) = 1,$$

where the sign $\sum$ denotes disjunction of events. Now the event which consists in the realization of an infinite number of the E_n's can be written as

$$\prod_{h=1}^{\infty}\left(\sum_{n=h}^{\infty} E_n\right)$$

where the sign $\prod$ denotes conjunction of events. Hence

$$\Pr\left(\prod_{h=1}^{\infty}\left(\sum_{n=h}^{\infty} E_n\right)\right) = \lim_{h\to\infty}\left(\sum_{n=h}^{\infty} E_n\right) = 1.$$

Thu. (1.5) is proved. The proof of (1.6) follows immediately from Lemma 5 and B. rel-Cantelli lemma.

PRINCETON UNIVERSITY

Reprinted from *Trans. Amer. Math. Soc.* 64 (1948), pp. 205–233
with permission from the American Mathematical Society.

ON THE MAXIMUM PARTIAL SUMS OF SEQUENCES OF INDEPENDENT RANDOM VARIABLES(1)

BY
KAI LAI CHUNG

1. Introduction. In this paper we deal with a sequence of independent random variables X_n, $n=1, 2, \cdots$. We write

$$S_n = \sum_{\nu=1}^{n} X_\nu, \tag{1}$$

$$S_n^* = \max_{1 \leq \nu \leq n} |S_\nu|. \tag{2}$$

Two types of fundamental limit theorems are known about S_n, the one clustering around the central limit theorem and the other the law of the iterated logarithm.

In 1945 Feller [12](2) called attention to the study of the behavior of S_n^*. Since then an important result has been obtained by Erdös and Kac [8], namely, the limiting distribution of S_n^* for sufficiently general sequences of X_n. This corresponds to the central limit theorem for S_n. Now under certain conditions when the distribution of S_n tends to the normal distribution, an estimate of the difference of the two distributions has been given by Liapounoff [17], Cramér [5], Berry [3] and Essen [9]. Cramér [6] and Feller [10] have also obtained more precise estimates for this difference for certain domains of variation of S_n, which proved essential to the general form of the law of the iterated logarithm. It is therefore of interest to make the same kind of investigations regarding S_n^*. The problem is more difficult, since we have as yet no standard tools as in the case of S_n. We shall prove in this direction, as consequences of a more general but less handy inequality (Lemma 7), two theorems corresponding to the two types of estimation mentioned above. In order to state them we introduce the following notations. Let $E(X)$ denote the mathematical expectation of X. We shall assume that for each X, the first moment is zero, and the third absolute moment is finite. Thus we can write

$$E(X_\nu) = 0; \tag{3}$$

$$E(X_\nu^2) = \sigma_\nu^2; \qquad s_n^2 = \sum_{\nu=1}^{n} \sigma_\nu^2; \tag{4}$$

Presented to the Society, September 4, 1947; received by the editors May 27, 1947.

(1) The present paper is the revised form of a Dissertation for the degree of Doctor of Philosophy accepted by Princeton University, 1947.

(2) Numbers in brackets refer to the references cited at the end of the paper.

$$E(|X_\nu|^3) = \gamma_\nu; \qquad \Gamma_n = \sum_{\nu=1}^{n} \gamma_\nu. \tag{5}$$

Naturally we assume that $s_n \to \infty$. We shall further make the following assumption:

$$\max_{1 \le \nu \le n} \gamma_\nu \sigma_\nu^{-2} = O(s_n^{1-\theta}) \tag{6}$$

where θ is a fixed but arbitrarily small positive number. Then we can prove the following two theorems.

THEOREM 1. *If c is a positive constant, then we have*

$$\begin{aligned} \Pr(S_n^* < cs_n) = {} & \frac{4}{\pi} \sum_{i=0}^{\infty} \frac{(-1)^i}{2i+1} \exp\left(-\frac{(2i+1)^2\pi^2}{8c^2}\right) \\ & + O\left(\left(\frac{\lg_2 s_n}{\lg s_n}\right)^{1/2}\right). \end{aligned} \tag{7}$$

THEOREM 2. If $g_n \downarrow 0$ *and*

$$g_n^{-1} = O((\lg_2 s_n)^{1/2}) \tag{8}$$

then we have(3)

$$\Pr(S_n^* < g_n s_n) = (1 + o(1)) \exp\left(-\frac{\pi^2}{8g_n^2}\right). \tag{9}$$

Theorem 2 is one of a number of possible statements; we give prominence to it here because it furnishes the means of proving the next group of theorems which we now consider.

We might attempt to extend the law of the iterated logarithm to S_n^*. This turns out to be illusory since the same law holds for S_n^* as for S_n. More precisely, if $\phi_n \uparrow \infty$, we have always ("i. o." standing for "infinitely often")

$$\Pr(S_n^* > \phi_n s_n \text{ i. o.}) = \Pr(S_n > \phi_n s_n \text{ i. o.}).$$

This is obvious since both S_n^* and $\phi_n s_n$ are monotone increasing functions of n. Hence in particular three of Feller's theorems [11] read as follows:

I. *If* $\sup |X_n| = O(s_n(\lg_2 s_n)^{-3/2})$ *and* $\phi_n^2 = 2 \lg_2 s_n + 3 \lg_3 s_n + 2 \lg_4 s_n + \cdots + 2 \lg_{p-1} s_n + (2+\delta) \lg_p s_n$ *then the probability*

$$\Pr(S_n^* > s_n \phi_n \text{ i. o.}) \tag{10}$$

(3) *Added in proof.* For the application of Theorem 2 in Lemma 9 it is important to notice that the constant in the $o(1)$ term in (9) depends only on the constants in the $O(1)$ terms in (6) and (8), and the θ in (6), but otherwise is independent of the random variables considered.

is equal to zero or one according as δ is positive or not.

II. *If $\phi_n \uparrow \infty$ and*

$$\sup |X_n| = O\left(\frac{s_n}{\phi_n^3}\right),$$

then (10) *is equal to zero or one according as the series*

$$\sum_n \frac{\sigma_n^2}{s_n^2} \phi_n e^{-(1/2)\phi_n^2}$$

is convergent or divergent.

III. *If $\phi(t) \uparrow \infty$ and*

$$\sup |X_n| = O\left(\frac{s_n}{\phi^3(s_n^2)}\right),$$

then $\Pr(S_n^* > s_n\phi(s_n^2)$ i. o.) *is equal to zero or one according as the integral*

$$\int^{\infty} \frac{1}{t} \phi(t) e^{-(1/2)\phi^2(t)} dt$$

is convergent or divergent.

These results give very precise upper bounds for S_n^*, with probability one. The question naturally arises as to the precise lower bounds for S_n^*. (We may mention that the analogous problem for S_n has been treated by Erdös and the author [4] and is radically different.) In this connection Erdös has communicated to the author the following result: there exist two constants $c_2 > c_1 > 0$ such that

$$\Pr\left(c_1 < \liminf \frac{S_n^*}{s_n(\lg_2 s_n)^{-1/2}} < c_2\right) = 1.$$

His method, of an elementary nature, does not seem capable of a sharper result. Using Theorem 2 stated above we can easily prove that

$$\Pr\left(\liminf \frac{S_n^*}{s_n(\lg_2 s_n)^{-1/2}} = 8^{-1/2}\pi\right) = 1.$$

This corresponds to Khintchine-Kolmogoroff's original form of the law of the iterated logarithm ([14] and [16]). However, we can go much further and prove the following theorems which are the exact counterparts of Feller's theorems cited above.

THEOREM 3. *Under the assumptions* (3) *to* (6), *if*

(11) $\phi_n^2 = \lg_2 s_n + 2 \lg_3 s_n + \lg_4 s_n + \cdots + \lg_{p-1} s_n + (1 + \delta) \lg_p s_n,$

then

$$\Pr(S_n^* < 8^{-1/2}\pi\phi_n^{-1}s_n \text{ i. o.}) \tag{12}$$

is equal to zero or one according as δ is positive or not.

THEOREM 4. *Under the same assumptions, if $\phi_n \uparrow \infty$, then* (12) *is equal to zero or one according as the series*

$$\sum \frac{\sigma_n^2}{s_n^2} \phi_n^2 e^{-\phi_n^2} \tag{13}$$

is convergent or divergent.

Theorem 3 is a particular case of Theorem 4.

THEOREM 5. *Under the same assumptions, if $\phi(s_n^2) \uparrow \infty$, then*

$$\Pr(S_n^* < 8^{-1/2}\pi\phi^{-1}(s_n^2)s_n \text{ i. o.}) \tag{14}$$

is equal to zero or one according as the integral

$$\int^\infty \frac{1}{t} \phi^2(t)e^{-\phi^2(t)}dt \tag{15}$$

is convergent or divergent.

The similarity between these theorems and Feller's is indeed striking. It should however be noted that the condition (6) is not the best possible, although it is weaker than those considered by Cramér [5]. That condition (6) can be trivially weakened will be apparent from the proof. But no complete settlement of the question seems in sight.

We outline the methods of proof as follows. We approximate the distribution on S_n^* by that of

$$\max_{1 \leq j \leq k} |S_{n_j}| \tag{16}$$

where k is an integer to be determined later and $0 < n_1 < \cdots < n_k = n$ is a suitably chosen sequence such that $s_{n_j}^2 \sim jk^{-1}s_n^2$.

In §2 we study the approximate distribution of (16). It is found to approach that of a k-dimensional normal distribution with a remainder we shall estimate. The treatment in Lemma 2(4), much to be preferred to the

(4) In the special case of equal components Bergström's result [2] seems to imply a better estimate than Lemma 2, replacing the factor 4^k by a fixed power of k. The improvement however is annulled by Lemma 3. It becomes essential in the problem of max S_ν, without the absolute value. We shall consider this elsewhere.

author's original proof using characteristic functions, is due to G. A. Hunt.

In §3 we estimate the difference between the distribution of S_n^* and that of (16). This is done by a substantial improvement of the method of Erdös and Kac (8), using sharper estimates resulting from the one-dimensional Berry-Esseen estimate. To obtain the approximate distribution of S_n^* it remains to evaluate the k-dimensional normal distribution obtained in §2. The problem appears to be one of multiple integrals but has not been worked out directly. Instead we use a quantitative refinement of the "invariance principle" of Erdös and Kac and study the simplest case of random walk. This latter problem, being almost classical, has been treated by many authors with different methods. However as we require not only the limiting distribution but also a remainder no reference seems available in the literature. We shall obtain the precise result by going back to a combinatorial formula due (apparently) to Bachelier [1]. After this we combine the results in §§2 and 3 to establish a theorem (Lemma 7) which includes Theorems 1 and 2 as particular cases.

In §4 we prove Theorems 3, 4 and 5. The proof of these theorems depends essentially on Theorem 2, which plays the role here as the theorem of Cramér-Feller does in the case of Feller's theorems cited above. Several arguments of Feller's are also used and the author's indebtedness to his previous work is considerable.

The author wishes to express his gratitude to Professor Cramér for his warm encouragement and valuable counsel. To Dr. Erdös, whose first result actually started the investigation, the author owes many heartfelt thanks for his sustained interest. To Mr. Hunt, who is responsible not only for Lemma 2 but for many corrections on the original manuscript, the author's gratitude is equally great.

2. An approximation theorem for a certain multi-dimensional distribution. We shall use $A_1, A_2, \cdots$ to denote absolute constants.

Let $n_1 < \cdots < n_k = n$ be a subsequence of $1, \cdots, n$ defined by the following:

$$s_{n_j}^2 \leqq jk^{-1}s_n^2 < s_{n_j+1}^2, \qquad j = 1, \cdots, k. \tag{17}$$

Then $(S_{n_1}, \cdots, S_{n_j})$ is a random point in j-dimensional space. Let its distribution function be

$$F_j(u_1, \cdots, u_j) = \Pr(S_{n_1} \leqq u_1, \cdots, S_{n_j} \leqq u_j). \tag{18}$$

Write also

$$F_j^*(x) = \Pr(S_{n_j} - S_{n_{j-1}} \leqq x), \qquad S_{n_0} = 0.$$

We put

$$B_j^2 = s_{n_j}^2 - s_{n_{j-1}}^2; \tag{19}$$

(20) $$M_n = \max_{1 \leq \nu \leq n} \gamma_\nu \sigma_\nu^{-2}.$$

LEMMA 1. *We have*

$$F_j^*(x) = \Phi_j^*(x) + R_j^*(x),$$

where $\Phi_j^*(x)$ *is the normal distribution function with mean* 0 *and variance* B_j^2, *and* $|R_j^*(x)| \leq A_1 M_n B_j^{-1}$.

This is a restatement of the Berry-Esseen theorem.

LEMMA 2. *Suppose that* (6) *holds and also that*

(21) $$\max_{1 \leq \nu \leq n} \sigma_\nu^2 = o(k^{-1} s_n).$$

Then we have

(22) $$|F_j(u_1, \cdots, u_j) - \Phi_j(u_1, \cdots, u_j)| \leq A_2 k^{1/2} 4^j M_n s_n^{-1},$$

where $\Phi_j(u_1, \cdots, u_j)$ *is the j-dimensional normal distribution function with the same moments of the first and second order as* $F_j(u_1, \cdots, u_j)$.

Proof. From (17), (19) and (21) it is easy to see that

(23) $$B_j \sim k^{-1/2} s_n.$$

Hence by Lemma 1, we have

(24) $$|R_j^*(x)| \leq A_2 k^{1/2} M_n s_n^{-1}.$$

For $j=1$, $R_1(x) = R_1^*(x)$; hence (22) is true for $j=1$. Now we use induction on j. Assume that

(25) $$|R_j(u_1, \cdots, u_j)| \leq A_2 k^{1/2} 4^j M_n s_n^{-1}.$$

We have, by the definition (18),

$$\begin{aligned} F_{j+1}(u_1, \cdots, u_{j+1}) &= \int_{-\infty}^{\infty} F_j(u_1, \cdots, u_{j-1}, \min(u_j, u_{j+1} - x)) dF_{j+1}^*(x) \\ &= \int_{-\infty}^{\infty} \{\Phi_j(u_1, \cdots, u_{j-1}, \min(u_j, u_{j+1} - x)) \\ &\quad + R_j(u_1, \cdots, u_{j-1}, \min(u_j, u_{j+1} - x))\} d[\Phi_{j+1}^*(x) + R_{j+1}^*(x)] \qquad (26) \\ &= \Phi_{j+1}(u_1, \cdots, u_{j+1}) + \int R_j d\Phi_{j+1}^* \\ &\quad + \int R_j d\Phi_{j+1}^* + \int \Phi_j dR_{j+1}^*. \end{aligned}$$

Evidently we have

$$\left|\int R_j d\Phi^*_{j+1}\right| \leqq \sup |R_j|,$$

$$\left|\int R_j dR^*_{j+1}\right| \leqq 2 \sup |R_j|.$$

Finally, using integration by parts, we have

$$\begin{aligned}
&\int_{-\infty}^{\infty} \Phi_j(u_1, \cdots, u_{j-1}, \min(u_j, u_{j+1} - x))dR^*_{j+1}(x) \\
&\quad = \int_{-\infty}^{u_{j+1}-u_j} \Phi_j(u_1, \cdots, u_j)dR^*_{j+1}(x) \\
&\qquad + \int_{u_{j+1}-u_j}^{\infty} \Phi_j(u_1, \cdots, u_j, u_{j+1} - x)dR^*_{j+1}(x) \\
&\quad = \Phi_j(u_1, \cdots, u_j)R^*_{j+1}(u_{j+1} - u_j) - \Phi_j(u_1, \cdots, u_j)R^*_{j+1}(u_{j+1} - u_j) \\
&\qquad + \int_{u_{j+1}-u_j}^{\infty} R^*_{j+1}(x)d\Phi_j(u_1, \cdots, u_{j-1}, u_{j+1} - x).
\end{aligned}$$

Hence the absolute value of the left-hand side is less than or equal to

$$\sup |R^*_{j+1}|.$$

Substituting these estimates into (22), we obtain

$$|F_{j+1} - \Phi_{j+1}| \leqq 3(\sup |R_j| + \sup |R^*_{j+1}|).$$

From (25) and (26), we have

$$\begin{aligned}
|F_{j+1} - \Phi_{j+1}| &\leqq 3A_2 k^{1/2} M_n s_n^{-1}(4^j + 1) \\
&\leqq A_2 4^{j+1} k^{1/2} M_n s_n^{-1}.
\end{aligned}$$

Thus the induction is complete.

Now we put, for non-negative u_j's,

$$F_0(u_1, \cdots, u_k) = \Pr(|S_{n_1}| \leqq s_n u_1, \cdots, |S_{n_k}| \leqq s_n u_k), \tag{27}$$

$$\begin{aligned}
\Phi_0(u_1, \cdots, u_k) = \frac{s_n^k}{(2\pi)^{k/2}B_1 \cdots B_k} \int_{-u_1}^{u_1} \cdots \int_{-u_k}^{u_k} \\
\cdot \exp\left(-\frac{1}{2}\sum_{j=1}^{k} \frac{s_n^2}{B_j^2}(x_j - x_{j-1})^2\right)dx_1 \cdots dx_k.
\end{aligned} \tag{28}$$

LEMMA 3. *Under the same assumptions as in Lemma 2, we have*

$$(29)\qquad |F_0 - \Phi_0| \leqq A_3(10)^k M_n s_n^{-1}.$$

Proof. Taking $j=k$ in (22), we have

$$(30)\qquad |F_k - \Phi_k| \leqq A_2 k^{1/2} 4^k M_n s_n^{-1} \leqq A_3 5^k M_n s_n^{-1}.$$

It is well known that we have

$$\begin{aligned} F_0(u_1, \cdots, u_k) = {} & F_k(s_n u_1, \cdots, s_n u_k) \\ & - F_k(-s_n u_1, s_n u_2, \cdots, s_n u_k) - \cdots \\ & - F_k(s_n u_1, \cdots, s_n u_{k-1}, -s_n u_k) \\ & + F_k(-s_n u_1, -s_n u_2, s_n u_3, \cdots, s_n u_k) + \cdots \\ & + (-1)^k F_k(-s_n u_1, \cdots, -s_n u_k); \end{aligned}$$

and a similar relation holds between Φ_k and Φ_0. Since there are 2^k terms on the right-hand side, (29) follows immediately from (30). It is not hard to obtain the explicit form of $\Phi_0(u_1, \cdots, u_k)$ in (28) by considering the covariance matrix.

3. **The distribution of the maximum partial sum.** Let c be a positive constant; g_n a monotone function of n; $\epsilon_n = o(1)$.

LEMMA 4. *Suppose that* (6) *and* (21) *are satisfied, and also that we have*

$$(31)\qquad \epsilon_n^2 g_n^2 = o(k^{-3/2} s_n^\theta),$$

$$(32)\qquad \sigma_n^2 = o((\epsilon_n^2 g_n^2 s_n^{3-\theta})^{2/3}).$$

Then we have

$$(33)\qquad \Pr(S_n^* < cg_n s_n) \geqq \Pr\left(\max_{1\leqq j\leqq k} |S_{n_j}| < (c - \epsilon_n) g_n s_n\right) - R_n$$

where

$$(34)\qquad R_n \leqq A_4(k^{-1/2}\epsilon_n^{-1} g_n^{-1} \exp(-4^{-1}\epsilon_n^2 g_n^2 k) + (\epsilon_n g_n s_n^\theta)^{-2/3}).$$

Proof. Write

$$P_n = \Pr(S_n^* < cg_n s_n),$$

$$W_r = \Pr(S_{r-1}^* < cg_n s_n, \ |S_r| \geqq cg_n s_n).$$

Then we have

$$(35)\qquad \sum_{r=1}^n W_r = \Pr(S_n^* \geqq cg_n s_n) = 1 - P_n \leqq 1.$$

Suppose that $n_j < \gamma \leqq n_{j+1}$. We have

$$
(36)\quad \begin{aligned} W_r = {} & \Pr(S^*_{r-1} < cg_n s_n,\ |S_r| \geqq cg_n s_n) \Pr(|S_{n_{j+1}} - S_r| \geqq \epsilon_n g_n s_n) \\ & + \Pr(S^*_{r-1} < cg_n s_n,\ |S_r| \geqq cg_n s_n,\ |S_{n_{j+1}} - S_r| < \epsilon_n g_n s_n). \end{aligned}
$$

Let $A > 0$ be an integer to be determined later. If $n_{j+1} - r \leqq A$, we have by the Tchebychef inequality,

$$
(37)\quad \Pr(|S_{n_{j+1}} - S_r| \geqq \epsilon_n g_n s_n) \leqq (s^2_{n_{j+1}} - s^2_{n_{j+1}-A+1})(\epsilon_n g_n s_n)^{-2}.
$$

If $n_{j+1} - r = B > A$, we have by the Berry-Esseen theorem,

$$
(38)\quad \Pr(|S_{n_{j+1}} - S_r| \geqq \epsilon_n g_n s_n) = \left(\frac{2}{\pi}\right)^{1/2} \int_v^\infty e^{-u^2/2} du + O(\rho)
$$

where

$$
v = \epsilon_n g_n s_n (s^2_{n_{j+1}} - s^2_{n_{j+1}-A+1})^{-1/2}
$$

and

$$
\rho = M_n \left(\sum_{\nu = n_{j+1}-A+1}^{n_{j+1}} \sigma_\nu^2 \right)^{-1/2}.
$$

Hence from (38), since $A < B \leqq n_{j+1} - n_j$, $s^2_{n_{j+1}} - s^2_{n_{j+1}-A+1} \leqq s^2_{n_{j+1}} - s^2_{n_j} = B^2_{j+1}$,

$$
(39)\quad \begin{aligned} & \Pr(|S_{n_{j+1}} - S_r| \geqq \epsilon_n g_n s_n) \\ & \leqq A_6 \left(\frac{B_{j+1}}{\epsilon_n g_n s_n} \exp\left(- \frac{\epsilon_n^2 g_n^2 s_n^2}{2B^2_{j+1}} \right) + \frac{M_n}{(s^2_{n_{j+1}} - s^2_{n_{j+1}-A+1})^{1/2}} \right). \end{aligned}
$$

We choose A such that

$$
\frac{M_n}{(s^2_{n_{j+1}} - s_{n_{j+1}-A+1})^{1/2}} \sim \frac{s^2_{n_{j+1}} - s^2_{n_{j+1}-A+1}}{\epsilon_n^2 g_n^2 s_n^2},
$$

that is,

$$
(s^2_{n_{j+1}} - s^2_{n_{j+1}-A+1})^{3/2} \sim \epsilon_n^2 g_n^2 s_n^2 M_n.
$$

Since $n_{j+1} - A + 1 \geqq n_j$ this is possible if, for example,

$$
\epsilon_n^2 g_n^2 s_n^2 M_n = o(B^3_{j+1}) = o(k^{-3/2} s_n^3)
$$

by (23), and also if

$$
\sigma_n^2 = o((\epsilon_n^2 g_n^2 s_n^2 M_n)^{2/3}).
$$

These are implied by the conditions (31) and (32), on account of (6). Hence we obtain from (37) and (39)

$$\Pr(|S_{n_{j+1}} - S_r| \geqq \epsilon_n g_n s_n) \leqq A_5\left(\frac{B_{j+1}}{\epsilon_n g_n s_n} \exp\left(-\frac{\epsilon_n^2 g_n^2 s_n^2}{2B_{j+1}^2}\right) + (\epsilon_n g_n s_n^{\theta})^{-2/3}\right).$$

Since $B_{j+1}^2 s_n^{-2} \sim k^{-1}$ by (23), we obtain

$$\Pr(|S_{n_{j+1}} - S_r| \geqq \epsilon_n g_n s_n) \leqq A_4\left(\frac{1}{k^{1/2}\epsilon_n g_n} \exp\left(-\frac{\epsilon_n^2 g_n^2 k}{4}\right) + (\epsilon_n g_n s_n^{\theta})^{-2/3}\right).$$

If we denote the maximum of the left-hand side of this inequality for all r by R_n, (36) becomes

$$(40) \qquad W_r \leqq R_n + \Pr(S_{r-1}^* < cg_n s_n, \ |S_r| \geqq cg_n s_n, \ |S_{n_{j+1}} - S_r| < \epsilon_n g_n s_n).$$

From (35) and (40), we obtain

$$\begin{aligned} &\Pr(S_n^* \geqq cg_n s_n) \\ &\leqq R_n + \sum_{j=0}^{k-1} \sum_{r=n_j+1}^{n_{j+1}} \Pr(S_{r-1}^* < cg_n s_n, \ |S_r| \geqq cg_n s_n, \ |S_{n_{j+1}} - S_r| < \epsilon_n g_n s_n) \\ &\leqq R_n + \Pr\left(\max_{1 \leqq j \leqq k} |S_{n_j}| \geqq (c - \epsilon_n) g_n s_n\right). \end{aligned}$$

This is equivalent to (33).

If in the function $F_0(u_1, \cdots, u_k)$ of (27) all the arguments are equal to u we shall use the shorter notation $F_{0k}(u)$; similarly for Φ_{0k}.

LEMMA 5. *Suppose that the condition* (6) *is satisfied, and also for a* $\Theta > 0$ *we have*

$$(41) \qquad \frac{8 \lg 10 \cdot \Theta}{\theta} \frac{\lg_2 s_n}{\lg s_n} \leqq \epsilon_n^2 g_n^2 = o\left(\frac{s_n^{\theta}}{(\lg s_n)^{3/2}}\right).$$

Then if we choose

$$(42) \qquad k \sim \frac{\theta \lg s_n}{2 \lg 10}$$

we have

$$(43) \qquad \Phi_{0k}((c - \epsilon_n) g_n) - L_n \leqq \Pr(S_n^* < cg_n s_n) \leqq \Phi_{0k}(cg_n) + L_n$$

where

$$(44) \qquad L_n = O((\lg s_n)^{-\Theta}).$$

Proof. From (6) it follows that

$$\sigma_n \leqq \gamma_n \sigma_n^{-2} = O(s_n^{1-\theta}).$$

Hence with the k in (42) condition (21) is satisfied. Further condition (32) in Lemma 4 is satisfied with $\epsilon_n^2 g_n^2$ satisfying (41). Condition (31) is clearly satisfied with (41) and the choice of k in (42). Hence both Lemma 3 and Lemma 4 are applicable. Taking all the u's in (29) to be $(c-\epsilon_n)g_n$ and recalling (27) we obtain

$$\begin{aligned} \Pr(S_n^* < cg_n s_n) &\geqq F_{0k}((c-\epsilon_n)g_n) - R_n \\ &\geqq \Phi_{0k}((c-\epsilon_n)g_n) - A_3(10)^k s_n^{-1} M_n - R_n. \end{aligned} \tag{45}$$

On the other hand we have

$$\Pr(S_n^* < cg_n s_n) \leqq F_{0k}(cg_n) \leqq \Phi_{0k}(cg_n) + A_3(10)^k s_n^{-1} M_n. \tag{46}$$

We find from (42) and (46)

$$(10)^k = O(s_n^{\theta/2}), \qquad (10)^k s_n^{-1} M_n = O(s_n^{-\theta/2}),$$

$$k^{-1/2}\epsilon_n^{-1} g_n^{-1} \exp(-4^{-1}\epsilon_n^2 g_n^2 k) = O((\lg s_n)^{-\Theta}),$$

$$(\epsilon_n g_n s_n^{\theta})^{-2/3} = O(s_n^{-\theta/2}).$$

Hence if we take

$$L_n = R_n + A_3(10)^k s_n^{-1} M_n = O((\lg s_n)^{-\Theta})$$

(45) and (46) imply (43).

LEMMA 6. *Suppose that for each* ν,

$$X_\nu = \begin{cases} +1 & \text{with probability } 1/2, \\ -1 & \text{with probability } 1/2. \end{cases} \tag{47}$$

Then if $g_n = o(n^{1/2})$, *we have*

$$\Pr(S_n^* < cg_n n^{1/2}) = T(cg_n) + O(g_n^{-1} n^{-1/2}) + O(n^{-1/2}) \tag{48}$$

where $T(x)$ *is the distribution function defined by*

$$T(x) = \frac{4}{\pi}\sum_{i=0}^{\infty} \frac{(-1)^i}{2i+1} \exp\left(-\frac{(2i+1)^2\pi^2}{8x^2}\right), \qquad x > 0. \tag{49}$$

Proof. Write, for integral a and b,

$$P(a) = \Pr(S_n = a, -b < S_\nu < b, \text{ for } 0 < \nu \leqq n).$$

By a formula due to Bachelier [1, pp. 252–253],

$$2^n P(a) = C_{n,(n+a)/2} + \sum_{1\leq i\leq (n+a)/2b} (-1)^i C_{n,(n+a)/2-ib} + \sum_{1\leq i\leq (n-a)/2b} (-1)^i C_{n,(n-a)/2-ib}$$

if n and a have the same parity, otherwise $P(a)=0$. Without loss of generality we may assume n to be even, b odd. Then

$$\sum_{-b<a<b,\, a\equiv 0 \pmod 2} C_{n,(n+a)/2-ib} = \sum_{(-b+1)/2\leq j\leq (b-1)/2} C_{n,n/2+j-ib}.$$

Write

$$P_i = P_{-i} = \sum_{(-b+1)/2\leq j\leq (b-1)/2} C_{n,n/2+j-ib}\frac{1}{2^n} = \sum_{n/2+1/2+(n^{1/2}/2)\zeta_{1i}\leq m\leq n/2-1/2+(n^{1/2}/2)\zeta_{2i}} C_{n,m};$$

where

$$\zeta_{1i} = -(2i+1)bn^{-1/2}, \qquad \zeta_{2i} = -(2i-1)bn^{-1/2}.$$

Finally we write

$$P = \sum_{-b<a<b} P(a). \tag{50}$$

From a formula of Uspensky [16, p. 129], noticing that the limits of the range of m are integers, we deduce easily that

$$\sum_{n/2+1/2+(n^{1/2}/2)\zeta_1\leq m\leq n/2-1/2+(n^{1/2}/2)\zeta_2} C_{n,m}\frac{1}{2^n} = \left(\frac{1}{2\pi}\right)^{1/2}\int_{\zeta_1}^{\zeta_2} e^{-u^2/2}du + O\left(\frac{1}{n}\right).$$

Hence

$$P_{-i} + P_i = \left(\frac{2}{\pi}\right)^{1/2}\int_{\zeta_{1i}}^{\zeta_{2i}} e^{-u^2/2}du + O\left(\frac{1}{n}\right). \tag{51}$$

Since $-b<a<b$,

$$\frac{1}{2^n}\left| \sum_{1\leq i\leq (n+a)/2b} (-1)^i C_{n,(n+a)/2-ib} + \sum_{1\leq i\leq (n-a)/2b} (-1)^i C_{n,(n-a)/2-ib} - \sum_{1\leq i\leq n/2b} (-1)^i [C_{n,(n+a)/2-ib} + C_{n,(n-a)/2-ib}]\right| \leq \frac{1}{2^n}.$$

Hence

$$\left| \sum_{-b<a<b,\, a\equiv 0 \pmod 2} P(a) - \sum_{1\leqq i\leqq n/2b} (-1)^i (P_i + P_{-i}) \right| \leqq \sum_{-b<a<b,\, a\equiv 0 \pmod 2} \frac{1}{2^n} \leqq \frac{b}{2^n}. \tag{52}$$

Therefore from (50) to (52),

$$\begin{aligned} P &= \left(\frac{1}{2\pi}\right)^{1/2} \int_{\zeta_{10}}^{\zeta_{20}} e^{-u^2/2} du + O\left(\frac{1}{n}\right) \\ &\quad + \sum_{1\leqq i\leqq n/2b} (-1)^i \left\{ \left(\frac{1}{2\pi}\right)^{1/2} \int_{\zeta_{1i}}^{\zeta_{2i}} e^{-u^2/2} du + O\left(\frac{1}{n}\right) \right\} + O\left(\frac{b}{2^n}\right) \\ &= \left(\frac{1}{2\pi}\right)^{1/2} \int_{\zeta_{10}}^{\zeta_{20}} e^{-u^2/2} du + \sum_{1\leqq i\leqq n/2b} (-1)^i \left(\frac{2}{\pi}\right)^{1/2} \int_{\zeta_{1i}}^{\zeta_{2i}} e^{-u^2/2} du \\ &\quad + O\left(\frac{1}{n}\right) + O\left(\frac{1}{b}\right) + O\left(\frac{b}{2^n}\right). \end{aligned} \tag{53}$$

Since the terms are alternating in sign and decreasing in absolute value, we have

$$\left| \sum_{i>n/2b} (-1)^i \left(\frac{2}{\pi}\right)^{1/2} \int_{\zeta_{1i}}^{\zeta_{2i}} e^{-u^2/2} du \right| \leqq \left(\frac{2}{\pi}\right)^{1/2} \int_{(n-b)n^{-1/2}}^{(n+b)n^{-1/2}} e^{-u^2/2} du = O(e^{-n/3}), \tag{54}$$

if $b=o(n)$.

Hence if $b=o(n)$, we obtain from (53) and (54),

$$\begin{aligned} P &= \left(\frac{1}{2\pi}\right)^{1/2} \int_{\zeta_{10}}^{\zeta_{20}} e^{-u^2/2} du + \sum_{i=1}^{\infty} (-1)^i \left(\frac{2}{\pi}\right)^{1/2} \int_{\zeta_{1i}}^{\zeta_{2i}} e^{-u^2/2} du + O\left(\frac{1}{b}\right) \\ &= \left(\frac{1}{2\pi}\right)^{1/2} \sum_{i=-\infty}^{\infty} (-1)^i \int_{(2i-1)bn^{-1/2}}^{(2i+1)bn^{-1/2}} e^{-u^2/2} du + O\left(\frac{1}{b}\right). \end{aligned} \tag{55}$$

We shall now construct a function $h(x)$ with period 2α as follows:

$$h(x) = \begin{cases} 1 & \text{if } 0 < x < \alpha/2, \\ -1 & \text{if } \alpha/2 < x < \alpha; \end{cases}$$

$$h(x) = h(-x); \qquad h(x) = h(x + 2\alpha).$$

It is easy to find that

$$h(x) = \frac{4}{\pi} \sum_{i=0}^{\infty} \frac{(-1)^i}{2i+1} \cos\left(\frac{2i+1}{\alpha} \pi x\right).$$

Taking α to be $2bn^{-1/2}$ in the above, we have

$$
\begin{aligned}
\left(\frac{1}{2\pi}\right)^{1/2} \sum_{i=-\infty}^{\infty} (-1)^i \int_{(2i-1)bn^{-1/2}}^{(2i+1)bn^{-1/2}} e^{-u^2/2} du &= \left(\frac{1}{2\pi}\right)^{1/2} \int_{-\infty}^{\infty} h(x) e^{-x^2/2} dx \\
&= \frac{4}{\pi} \sum_{i=0}^{\infty} \frac{(-1)^i}{2i+1} \left(\frac{1}{2\pi}\right)^{1/2} \int_{-\infty}^{\infty} e^{-x^2/2} \cos\left(\frac{(2i+1)\pi n^{1/2}}{2b} x\right) dx \qquad (56) \\
&= \frac{4}{\pi} \sum_{i=0}^{\infty} \frac{(-1)^i}{2i+1} \exp\left(-\frac{(2i+1)^2\pi^2 n}{8b^2}\right)
\end{aligned}
$$

since

$$\left(\frac{1}{2\pi}\right)^{1/2} \int_{-\infty}^{\infty} e^{-x^2/2} \cos tx dx = e^{-t^2/2}.$$

Therefore from (55) and (56) we obtain

$$P = T(bn^{-1/2}) + O(b^{-1}). \tag{57}$$

Since by assumption $g_n = o(n^{1/2})$, $cg_n n^{1/2} = o(n)$; taking b successively to be the nearest odd integers to $cg_n n^{1/2}$ in (54) and observing that $T(bn^{-1/2}) - T(cg_n) = O(n^{-1/2})$ we obtain (48).

Lemma 7. *Returning to the general case, we have, if* (6) *and* (41) *are satisfied,*

$$T((c - \epsilon_n)g_n) - H_n \leqq \Pr(S_n^* < cg_n s_n) \leqq T((c + \epsilon_n)g_n) + H_n; \tag{58}$$

where $T(x)$ is defined in (49) *and*

$$H_n = O((\lg s_n)^{-\Theta} + g_n^{-1} s_n^{-1}). \tag{59}$$

Proof. For the special case (47), we have according to the general notation (4) and (5),

$$\sigma_m^2 = 1, \qquad \gamma_m = 1, \qquad s_m^2 = m, \qquad M_m = 1.$$

Condition (6) is satisfied with $\theta = 1$. Hence by Lemma 5, if

$$\frac{8 \lg 10 \cdot \Theta \lg_2 m^{1/2}}{\theta \lg m^{1/2}} \leqq \epsilon_m'^2 g_m'^2 = o\left(\frac{m^{1/2}}{(\lg m)^{3/2}}\right)$$

we have from (43),

$$\Phi_{1k}((c - \epsilon_m')g_m') - L_m \leqq \Pr(S_m^* < cg_m' m^{1/2}) \leqq \Phi_{1k}(cg_m') + L_m, \tag{60}$$

where, by (42), k is given by

$$k \sim \frac{\theta \lg m}{4 \lg 10},$$

and where Φ_{1k} is obtained from Φ_{0k} in (28) after we replace s_n^2 by m and B_j by B_j' defined according to (17) and (19) by

$$m_j \leqq jk^{-1}m < m_j + 1, \; B_j'^2 = m_j - m_{j-1},$$

and where

$$L_m = O((\lg m)^{-\Theta}).$$

On the other hand, by Lemma 6, we have for the special case in question

$$\text{(61)} \qquad \Pr(S_m^* < cg_m' m^{1/2}) = T(cg_m') + O(g_m'^{-1} m^{-1/2}) + O(m^{-1/2}).$$

Substituting from (58) into (60) we obtain

$$\text{(62)} \qquad \Phi_{1k}(cg_m') - L_m \leqq T(cg_m') + O(g_m'^{-1} m^{-1/2} + m^{-1/2}) \leqq \Phi_{1k}(cg_m') + L_m.$$

From (62) we deduce

$$\text{(63)} \qquad \Phi_{1k}((c - \epsilon_m')g_m') \geqq T((c - \epsilon_m')g_m') + O(g_m'^{-1} m^{-1/2} + m^{-1/2}) + O(L_m);$$

$$\text{(64)} \qquad \Phi_{1k}(cg_m') \leqq T((c + \epsilon_m')g_m') + O(g_m'^{-1} m^{-1/2} + m^{-1/2}) + O(L_m).$$

Now putting

$$m = [s_n^2], \qquad \epsilon_m' = \epsilon_n, \qquad g_m' = g_n,$$

we obtain from (63) and (64) the following: if

$$\frac{8 \lg 10 \cdot \Theta}{\theta} \, \frac{\lg_2 s_n}{\lg s_n} \leqq \epsilon_n^2 g_n^2 = o\left(\frac{s_n^\theta}{(\lg s_n)^{3/2}}\right)$$

then we have

$$\text{(65)} \qquad \begin{aligned} T((c - \epsilon_n)g_n) - K_n \leqq \Phi_{1k}((c - \epsilon_n)g_n) &\leqq \Phi_{1k}(cg_n) \\ &\leqq T((c + \epsilon_n)g_n) + K_n, \end{aligned}$$

where

$$\text{(42 bis)} \qquad k \sim \frac{\theta \lg s_n}{2 \lg 10},$$

$$\text{(66)} \qquad K_n = O((\lg s_n)^{-\Theta} + g_n^{-1} s_n^{-1}).$$

Writing $\lambda_j = s_n B_j^{-1}$, $\lambda_j' = [s_n^2]^{1/2} B_j'^{-1}$ we have from (28)

$$\Phi_0(u_1, \cdots, u_k)$$

$$= \frac{\lambda_1 \cdots \lambda_k}{(2\pi)^{k/2}} \int_{-u_1}^{u_1} \cdots \int_{-u_k}^{u_k} \exp\left(-\frac{1}{2} \sum_{j=1}^{k} \lambda_j^2 (x_j - x_{j+1})^2\right) dx_1 \cdots dx_k.$$

It is easy to verify that

$$\left|\frac{\partial \Phi_0}{\partial \lambda_j}\right| \leqq \frac{3}{2\lambda_j}\,.$$

Hence

$$(67)\qquad | \Phi_{0k} - \Phi_{1k} | \leqq \frac{3}{2} \sum_{j=1}^{k} \frac{1}{\lambda_j} | \lambda_j - \lambda_j' |.$$

Since $\sigma_n = O(s_n^{1-\theta})$, $B_j^2 = s_n^2 k^{-1} + O(s_n^{2-2\theta})$,

$$\lambda_j^2 = \frac{s_n^2}{B_j^2} = k\left(1 + O\left(\frac{k}{s_n^{2\theta}}\right)\right).$$

The same holds for $\lambda_j'^2$. Thus $|\lambda_j - \lambda_j'| = O(k^2 s_n^{-2\theta})$; and by (42 bis) and (67),

$$(68)\qquad | \Phi_{0k} - \Phi_{1k} | = O(s_n^{-\theta}).$$

Therefore from (65) we obtain

$$(69)\qquad \begin{aligned} T((c - \epsilon_n)g_n) - J_n \leqq \Phi_{0k}((c - \epsilon_n)g_n) &\leqq \Phi_{0k}(cg_n) \\ &\leqq T((c + \epsilon_n)g_n) + J_n, \end{aligned}$$

where k is given by (42 bis) and from (66) and (68) we have

$$(70)\qquad J_n = O(R_n) + O(s_n^{-\theta}) = O((\lg s_n)^{-\Theta} + g_n^{-1} s_n^{-1}).$$

Using (69) in (43), Lemma 5, we obtain

$$T((c - \epsilon_n)g_n) - H_n \leqq \Pr (S_n^* < cg_n s_n) \leqq T((c + \epsilon_n)g_n) + H_n$$

where $H_n = O(J_n) + O(L_n)$. Hence by (44) and (70) we have established (58) and (59).

Proof of Theorem 1. Taking $g_n = 1$ in Lemma 5, we get

$$(71)\qquad T(c - \epsilon_n) - H_n \leqq \Pr(S_n^* < cs_n) \leqq T(c + \epsilon_n) + H_n,$$

where

$$H_n = O((\lg s_n)^{-\Theta}).$$

Now we have, by the mean-value theorem,

$$\exp\left(\frac{-(2i+1)^2\pi^2}{8(c+\epsilon_n)^2}\right) - \exp\left(\frac{-(2i+1)^2\pi^2}{8c^2}\right) \leqq \frac{(2i+1)^2\pi^2}{4c^3} \exp\left(-\frac{(2i+1)^2\pi^2}{8(c+\epsilon_n)^2}\right)\epsilon_n.$$

Hence

$$T(c+\epsilon_n)-T(c)\leqq \frac{4}{\pi}\sum_{i=0}^{\infty}\frac{(2i+1)\pi^2}{4c^3}\exp\left(-\frac{(2i+1)^2\pi^2}{8(c+\epsilon_n)^2}\right)\epsilon_n = O(\epsilon_n).$$

Thus (71) becomes

$$\Pr(S_n^* < cs_n) = T(c) + O(\epsilon_n) + O((\lg s_n)^{-\Theta}). \tag{72}$$

Choosing, for example, $\Theta=1$ and

$$\epsilon_n^2 = \frac{8 \lg 10}{\theta}\, \frac{\lg_2 s_n}{\lg s_n}$$

which is permissible by (41), we obtain (7) from (72).

Proof of Theorem 2. We have

$$\frac{4}{\pi}\exp\left(-\frac{\pi^2}{8x^2}\right)-\frac{4}{3\pi}\exp\left(-\frac{9\pi^2}{8x^2}\right)\leqq T(x)\leqq \frac{4}{\pi}\exp\left(-\frac{\pi^2}{8x^2}\right).$$

Since $\epsilon_n \downarrow 0$, we have if $\epsilon_n<4^{-1}$,

$$\begin{aligned} T((1+\epsilon_n)g_n) &\leqq \frac{4}{\pi}\exp\left(-\frac{\pi^2}{8g_n^2}(1+\epsilon_n)^{-2}\right)\\ &\leqq \frac{4}{\pi}\exp\left(-\frac{\pi^2}{8g_n^2}(1-2\epsilon_n)\right) \qquad (73)\\ &= \frac{4}{\pi}\exp\left(-\frac{\pi^2}{8g_n^2}\right)\exp\left(\frac{\pi^2\epsilon_n}{4g_n^2}\right), \end{aligned}$$

$$\begin{aligned} T((1-\epsilon_n)g_n) &\geqq \frac{4}{\pi}\exp\left(-\frac{\pi^2}{8g_n^2}(1-\epsilon_n)^{-2}\right)-\frac{4}{3\pi}\exp\left(-\frac{9\pi^2}{8g_n^2}\right)\\ &\geqq \frac{4}{\pi}\exp\left(-\frac{\pi^2}{8g_n^2}(1+4\epsilon_n)\right)-\frac{4}{3\pi}\exp\left(-\frac{9\pi^2}{8g_n^2}\right) \qquad (74)\\ &\geqq \frac{4}{\pi}\exp\left(-\frac{\pi^2}{8g_n^2}\right)\exp\left(-\frac{\pi^2\epsilon_n}{2g_n^2}\right)-\frac{4}{3\pi}\exp\left(-\frac{9\pi^2}{8g_n^2}\right). \end{aligned}$$

Choosing

$$\epsilon_n^2 = \frac{8\lg 10\cdot\Theta}{\theta}\,\frac{\lg_2 s_n}{g_n^2 \lg s_n}$$

then (41) is satisfied, and from (8),

$$\frac{\epsilon_n^2}{g_n^4} = \frac{8\lg 10\cdot\Theta}{\theta}\,\frac{\lg_2 s_n}{g_n^6\lg s_n} = o(1).$$

Hence we have

$$\exp\left(\frac{\pi^2\epsilon_n}{4g_n^2}\right) = 1 + o(1),$$

$$\exp\left(-\frac{\pi^2\epsilon_n}{2g_n^2}\right) = 1 + o(1).$$

Since $g_n \downarrow 0$, we have

$$\exp\left(-\frac{9\pi^2}{8g_n^2}\right) = o\left(\exp\left(-\frac{\pi^2}{8g_n^2}\right)\right).$$

Thus from (73) and (74), we obtain

$$\frac{4}{\pi}(1 + o(1)) \exp\left(-\frac{\pi^2}{8g_n^2}\right) \leqq T((1-\epsilon_n)g_n) \leqq T((1+\epsilon_n)g_n)$$
$$\leqq \frac{4}{\pi}(1 + o(1)) \exp\left(-\frac{\pi^2}{8g_n^2}\right).$$

Therefore (58) becomes

$$\Pr(S_n^* < g_n s_n) = \frac{4}{\pi}(1 + o(1)) \exp\left(-\frac{\pi^2}{8g_n^2}\right) + O((\lg s_n)^{-\Theta}).$$

Since we may choose Θ arbitrarily large, (9) follows on account of (8).

4. **Some strong limit theorems.** Since we shall deal with indices n, ν, k and so on, which ultimately tend to infinity, we shall often omit mention of this proviso. Thus, sometimes our statements are true only when the appropriate index is sufficiently large.

The condition (6) is assumed in this section. From (6) it follows:

$$\sigma_n = O(s_n^{1-\theta}), \qquad \theta > 0. \tag{75}$$

Let $\psi_n \uparrow \infty$, and

$$\psi_n = O((\lg_2 s_n)^{1/2}). \tag{76}$$

Taking $g_n = \psi_n^{-1}$ in Theorem 2, we have

$$A_6 e^{-\psi_n^2} \leqq \Pr(S_n^* < 8^{-1/2}\pi s_n \psi_n^{-1}) \leqq A_7 e^{-\psi_n^2}. \tag{9 bis}$$

We shall construct a subsequence $\{n_k\}$, $k = 1, 2, \cdots$, as follows. Take $a > 0$. Put $n_1 = 1$. Suppose that n_k is defined already, then since $s_n \uparrow$, there is a unique n_{k+1} such that

$$s_{n_{k+1}-1} \leqq s_{n_k}(1 + a\psi_{n_k}^{-2}) \leqq s_{n_{k+1}}.$$

Hence (for k sufficiently large)

$$s_{n_{k+1}-1}^2 \leqq s_{n_k}^2(1 + 3a\psi_{n_k}^{-2}).$$

By virtue of (75) and (76), we have

$$s_{n_{k+1}}^2 \leqq s_{n_k}^2 + 3as_{n_k}^2\psi_{n_k}^{-2} + \sigma_{n_{k+1}}^2 \leqq s_{n_k}^2 + 4as_{n_{k+1}}^2\psi_{n_k}^{-2};$$

$$s_{n_{k+1}}^2 \leqq s_{n_k}^2(1 - 4a\psi_{n_k}^{-2})^{-1}.$$

Thus there exists $b>a$ such that

$$s_{n_k}(1 + a\psi_{n_k}^{-2}) \leqq s_{n_{k+1}} \leqq s_{n_k}(1 + b\psi_{n_k}^{-2}). \tag{77}$$

For simplicity we shall write k' for n_k, s_k' for s_{n_k}, ψ_k' for ψ_{n_k}, and so on.

LEMMA 8. *Suppose that $\psi_n \uparrow \infty$ and (76) holds. Let $\{n_k\}$ be any sequence satisfying (77). Then if*

$$\sum_k e^{-\psi_{n_k}^2} < \infty \tag{78}$$

we have

$$\Pr(S_n^* < 8^{-1/2}\pi s_n\psi_n^{-1} \text{ i. o.}) = 0. \tag{79}$$

Proof. From (9 bis) we have

$$\Pr\left(S_{k'}^* < \frac{\pi}{8^{1/2}} \frac{s_k'}{(\psi_k'^2 - 3b)^{1/2}}\right) \leqq A_7 e^{-(\psi_k'^2 - 3b)}.$$

Hence by (78)

$$\sum_{k=1}^{\infty} \Pr\left(S_{k'}^* < \frac{\pi}{8^{1/2}} \frac{s_k'}{(\psi_k'^2 - 3b)^{1/2}}\right) < \infty.$$

By the lemma of Borel-Cantelli (see, for example [13, pp. 26–27]), we conclude that

$$\Pr\left(S_{k'}^* < \frac{\pi}{8^{1/2}} \frac{s_k'}{(\psi_k'^2 - 3b)^{1/2}} \text{ i. o.}\right) = 0,$$

that is,

$$\Pr\left(S_{k'}^* \geqq \frac{\pi}{8^{1/2}} \frac{s_k'}{(\psi_k'^2 - 3b)^{1/2}} \text{ for all sufficiently large } k\right) = 1. \tag{80}$$

Now suppose that $n_k < n \leqq n_{k+1}$. Then if

$$S_{k'}^* \geqq \frac{\pi}{8^{1/2}} \frac{s_k'}{(\psi_k'^2 - 3b)^{1/2}}$$

we have by (77)

$$S_n^* \geqq S_{k'}^* \geqq \frac{\pi}{8^{1/2}} \frac{s'_{k+1}}{(\psi_k'^2 - 3b)^{1/2}} \frac{s'_k}{s'_{k+1}}$$
$$\geqq \frac{\pi}{8^{1/2}} s'_{k+1}(\psi_k'^2 - 3b)^{-1/2}(1 + b\psi_k'^{-2})^{-1}.$$

If k is sufficiently large, we have

$$S_n^* \geqq \frac{\pi}{8^{1/2}} \frac{s'_{k+1}}{\psi'_k} \geqq \frac{\pi}{8^{1/2}} \frac{s_n}{\psi_n}.$$

Thus (80) entails

$$\Pr\left(S_n^* > \frac{\pi}{8^{1/2}} \frac{s_n}{\psi_n} \text{ for all sufficiently large } n\right) = 1.$$

This is equivalent to (79).

LEMMA 9. *Suppose that $\psi_n \uparrow \infty$ and* (76) *holds. Let $\{n_k\}$ be any sequence satisfying* (77). *Then if*

$$\sum_k e^{-\psi_{n_k}^2} = \infty \tag{81}$$

we have

$$\Pr\left(S_n^* < \frac{\pi}{8^{1/2}} \frac{s_n}{\psi_n} \text{ i. o.}\right) = 1. \tag{82}$$

Proof. By (77), given $s'_{k_{\nu-1}}$, there is a unique ν such that

$$s'_{k_\nu} \leqq s'_{k_{\nu-1}} \psi'^3_{k_{\nu-1}} < s'_{k_\nu+1}. \tag{83}$$

From (9 bis) we have, if $c > 1/8$ is any constant,

$$\Pr\left(S_{k'}^* < \frac{\pi}{8^{1/2}} \frac{s'_k}{(\psi_k'^2 + 8C)^{1/2}}\right) \geqq A_6 e^{-(\psi_k'^2 + 8C)}.$$

Hence by (81) we have

$$\sum_{\nu=1}^{\infty} \sum_{k=k_\nu}^{k_{\nu+1}-1} \Pr\left(S_{k'}^* < \frac{\pi}{8^{1/2}} \frac{s'_k}{(\psi_k'^2 + 8C)^{1/2}}\right) = \infty. \tag{84}$$

Let $\{\nu(r)\}$, $r = 1, 2, \cdots$, denote the subsequence of $\nu = 1, 2, \cdots$ for which

$$\psi^2(k'_{\nu(r)+1}) > \psi^2(k'_{\nu(r)-1}) + 1. \tag{85}$$

Then we have

$$\Pr\left(S_{k'}^{*} < \frac{\pi}{8^{1/2}} \frac{s_k'}{(\psi_k' + 8C)^{1/2}}\right) \leqq A_7 e^{-(\psi_k'^2+8C)} \leqq A_8 \psi^{-2}(k') e^{-\psi^2(k')/2}.$$

From (83),

$$\psi_{k_\nu}'^3 \geqq \frac{s_{k_{\nu+1}}'}{s_{k_\nu}'} \geqq \prod_{k=k_\nu}^{k_{\nu+1}-1}\left(1+\frac{a}{\psi'^2}\right) \geqq a \sum_{k=k_\nu}^{k_{\nu+1}-1} \frac{1}{\psi'^2}.$$

Hence

$$\sum_{k=k_{\nu(r)}}^{k_{\nu(r)+1}-1} \Pr\left(S_{k'}^{*} < \frac{\pi}{8^{1/2}} \frac{s_k'}{(\psi_k'^2 + 8C)^{1/2}}\right) \leqq A_8 e^{-\psi^2(k_{\nu(r)}')/2} \sum_{k=k_{\nu(r)}}^{k_{\nu(r)+1}-1} \frac{}{\psi_k'^2}$$

$$\leqq \frac{A_8}{a} \psi_{k_{\nu(r)}}'^3 e^{-\psi^2(k_{\nu(r)}')/2} \leqq A_9 e^{-\psi^2(k_{\nu(r)}')/4}.$$

Thus by (85)

$$(86) \qquad \sum_r \sum_{k=k_{\nu(r)}}^{k_{\nu(r)+1}-1} \Pr\left(S_{k'}^{*} < \frac{\pi}{8^{1/2}} \frac{s_k'}{(\psi_k'^2 + 8C)^{1/2}}\right) \leqq A_9 \sum_r e^{-\psi^2(k_{\nu(r)}')/4} < \infty.$$

From (82) and (86) we conclude that if we delete the values of ν equal to $\nu(r)$, $r=1, 2, \cdots$, in (84) the remaining series is still divergent. Without loss of generality we may then assume that, for some fixed ν_0,

$$(87) \qquad \sum_{\nu=1,\nu\equiv\nu_0 \pmod 2}^{\infty}{}' \sum_{k=k_\nu}^{k_{\nu+1}-1} \Pr\left(S_{k'}^{*} < \frac{\pi}{8^{1/2}} \frac{s_k'}{(\psi_k'^2 + 8C)^{1/2}}\right) = \infty,$$

where the prime after the summation indicates the omission of the values $\nu(r)$.

Denote by:

E_μ the event

$$S_\mu^{*} < \frac{\pi}{8^{1/2}} \frac{s_\mu}{\psi_\mu},$$

$E_{\nu-1}'$ the event

$$S_{k_{\nu-1}'}^{*} \leqq C s_{k_{\nu-1}'},$$

$E_{\nu-1,\mu}$ the event

$$\max_{k_{\nu-1}' < \rho \leqq \mu} \left| S_\rho - S_{k_{\nu-1}'} \right| < \frac{\pi}{8^{1/2}} \frac{s_\mu}{(\psi_\mu^2 + 8C)^{1/2}}, \qquad k_\nu' \leqq \mu \leqq k_{\nu+1}'.$$

If $\nu \neq \nu(r)$, then from (83) and (77) we have, since $\psi_{k_\nu}'^2 \leqq \psi_{k_{\nu-1}}'^2 + 1$, $\psi_{k_\nu}'^3 \leqq 2^{1/2}\psi_{k_{\nu-1}}'^3$ for large ν,

$$(88) \qquad s_{k_{\nu-1}}' < \psi_{k_{\nu-1}}'^{-3} s_{k_\nu}' (1 + b\psi_{k_\nu}'^{-2}) \leqq s_{k_\nu}' 2^{1/2} \psi_{k_\nu}'^{-3} (1 + b\psi_{k_\nu}'^{-2}) \leqq 2 s_{k_\nu}' \psi_{k_\nu}'^{-3}.$$

Then if we have the conjunction $E'_{\nu-1}E_{\nu-1,\mu}$, we have by (88)

$$S^*_\mu < \frac{\pi}{8^{1/2}} \frac{s_\mu}{(\psi_\mu^2 + 8C)^{1/2}} + Cs'_{k_{\nu-1}} < \frac{\pi}{8^{1/2}}\left(\frac{s_\mu}{(\psi_\mu^2 + 8C)^{1/2}} + \frac{2Cs_\mu}{\psi_\mu^3}\right)$$
$$< \frac{\pi}{8^{1/2}} \frac{s_\mu}{\psi_\mu}\left(\frac{1}{(1 + 8C\psi_\mu^{-2})^{1/2}} + \frac{2C}{\psi_\mu^2}\right) < \frac{\pi}{8^{1/2}} \frac{s_\mu}{\psi_\mu}.$$

Therefore if $\nu \neq \nu(r)$, the conjunction $E'_{\nu-1}E_{\nu-1,\mu}$ implies E_μ. Writing

$$F_\nu = \sum_{\mu=k'_\nu}^{k'_{\nu+1}-1} E_\mu, \qquad F'_\nu = \sum_{\mu=k'_\nu}^{k'_{\nu+1}-1} E_{\nu-1,\mu},$$

we have, if $\nu \neq \nu(r)$, $E'_{\nu-1}F'_\nu$ implies F_ν, hence

$$\sum_{\nu=\nu_1}^{\infty}{}' E'_{\nu-1}F'_\nu \quad \text{implies} \quad \sum_{\nu=\nu_1}^{\infty}{}' F_\nu,$$

$$\Pr\left(\sum_{\nu=\nu_1}^{\infty}{}' E'_{\nu-1}F'_\nu\right) \leq \Pr\left(\sum_{\nu=\nu_1}^{\infty}{}' F_\nu\right). \tag{89}$$

The events $F'_\nu, F'_{\nu+2}, F'_{\nu+4}, \cdots$ are independent and F'_j for $j \geq \nu$ is independent of $E'_{\nu-1}$. By the Kolmogoroff inequality [15] we have

$$\Pr(E'_{\nu-1}) \geq 1 - 1/C^2. \tag{90}$$

We obtain, by an obvious argument and (90), for all $\nu_1 \geq \nu_0$,

$$\begin{aligned}\Pr\left(\sum_{\nu=\nu_1}^{\infty}{}' E'_{\nu-1}F'_\nu\right) &= \Pr\left(\sum_{\nu=\nu_1}^{\infty}{}' (E'_{\nu-1}F'_\nu - E'_{\nu-1}F'_\nu \sum_{j=\nu+1}^{\infty} E'_{j-1}F'_j\right) \\ &\geq \Pr\left(\sum_{\nu=\nu_1}^{\infty}{}' E'_{\nu-1}\left(F'_\nu - F'_\nu \sum_{j=\nu+1}^{\infty}{}' F'_j\right)\right) \\ &= \sum_{\nu=\nu_1}^{\infty}{}' \Pr(E'_{\nu-1}) \Pr\left(F'_\nu - F'_\nu \sum_{j=\nu+1}^{\infty}{}' F'_j\right) \\ &\geq \left(1 - \frac{1}{C^2}\right) \Pr\left(\sum_{\nu=\nu_1}^{\infty}{}' F'_\nu\right).\end{aligned} \tag{91}$$

Hence by (89) and (91) we have

$$\begin{aligned}\Pr\left(\sum_{\nu=\nu_1}^{\infty}{}' F_\nu\right) &\geq \left(1 - \frac{1}{C^2}\right) \Pr\left(\sum_{\nu=\nu_1}^{\infty}{}' F'_\nu\right) \\ &\geq \left(1 - \frac{1}{C^2}\right) \Pr\left(\sum_{\nu=\nu_1, \nu\equiv\nu_0 \pmod 2}^{\infty}{}' F'_\nu\right).\end{aligned} \tag{92}$$

Since the events $F'_{\nu_0}, F'_{\nu_0+2}, F'_{\nu_0+4}, \cdots$ are independent, by the lemma of

Borel-Cantelli we know that

$$\Pr\Big(\sum_{\nu=\nu_1,\nu\equiv\nu_0 \pmod 2}^{\infty}{}' F_\nu'\Big) = 1$$

if and only if

$$\sum_{\nu=1,\nu\equiv\nu_0 \pmod 2}^{\infty}{}' \Pr(F_\nu') = \infty. \tag{93}$$

If we can prove (93), then from this remark and (91) we shall have for all $\nu_1 \geqq \nu_0$, hence in fact for all ν_1,

$$\Pr\Big(\sum_{\nu=\nu_1}^{\infty}{}' F_\nu\Big) \geqq 1 - \frac{1}{C^2},$$

a fortiori, for all n_1,

$$\Pr\Big(\sum_{n=n_1}^{\infty} E_n\Big) \geqq 1 - \frac{1}{C^2}\cdot$$

Since we may choose C arbitrarily large while the left-hand side does not depend on C we shall have proved for all n_1, $\Pr(\sum_{n=n_1}^{\infty} E_n) = 1$, which is equivalent to (82).

Hence to prove (82) it is sufficient to prove (93). By definition this is equivalent to

$$\sum_{\nu=1,\nu\equiv\nu_0 \pmod 2}^{\infty}{}' \Pr\Big(\sum_{\mu=k_\nu'}^{k_{\nu+1}'-1} E_{\nu-1,\mu}\Big) = \infty. \tag{94}$$

Comparing (94) and (87) we see that in order to prove (94) it is sufficient to prove that for $\nu \neq \nu(r)$, there exists a constant $A_{10} > 0$ such that for all sufficiently large ν, the following shall hold:

$$\Pr\Big(\sum_{\mu=k_\nu'}^{k_{\nu+1}'-1} E_{\nu-1,\mu}\Big) \geqq A_{10} \sum_{k=k_\nu}^{k_{\nu+1}-1} \Pr\Big(S_{k'}^* < \frac{\pi}{8^{1/2}} \frac{s_k'}{(\psi_k'^2 + 8C)^{1/2}}\Big). \tag{95}$$

We have for any integer $N > 0$,

$$\begin{aligned} \Pr\Big(\sum_{\mu=k_\nu'}^{k_{\nu+1}'-1} E_{\nu-1,\mu}\Big) &\geqq \Pr\Big(\sum_{k=k_\nu}^{k_{\nu+1}-1} E_{\nu-1,k'}\Big) \\ &\geqq \frac{1}{N} \sum_{k=k_\nu}^{k_{\nu+1}-1} \Pr\Big(E_{\nu-1,k'} - E_{\nu-1,k'} \sum_{j=k+N}^{k_{\nu+1}} E_{\nu-1,j'}\Big). \end{aligned} \tag{96}$$

Now we see easily that $E_{\nu-1,k'}E_{\nu-1,j'}$ implies $E_{k,j}'$ where $E_{k,j}'$ denotes the event

$$\max_{n_k < i \leqq n_j} |S_i - S_{n_k}| < \frac{\pi}{8^{1/2}}\left(\frac{s_{n_k}}{(\psi^2_{n_k} + 8C)^{1/2}} + \frac{s_{n_j}}{(\psi^2_{n_j} + 8C)^{1/2}}\right).$$

Since $E_{\nu-1,k'}$ and $E'_{k,j}$ are independent, we have

$$\Pr(E_{\nu-1,k'}E_{\nu-1,j'}) \leqq \Pr(E_{\nu-1,k'}) \Pr(E'_{k,j}).$$

If we can prove that, for a suitable N,

$$\sum_{j=k+N}^{k_{\nu+1}} \Pr(E'_{k,j}) < \frac{1}{2}, \tag{97}$$

then from (96)

$$\begin{aligned}\Pr\left(\sum_{\mu=k'_\nu}^{k'_{\nu+1}} E_{\nu-1,\mu}\right) &\geqq \frac{1}{N}\sum_{k=k_\nu}^{k_{\nu+1}} \Pr\left(E_{\nu-1,k'} - \sum_{j=k+N}^{k_{\nu+1}} E_{\nu-1,k'}E_{\nu-1,j'}\right)\\ &\geqq \frac{1}{N}\sum_{k=k_\nu}^{k_{\nu+1}} \Pr(E_{\nu-1,k'})\left(1 - \sum_{j=k+N}^{k_{\nu+1}} \Pr(E'_{k,j})\right)\\ &\geqq \frac{1}{2N}\sum_{k=k_\nu}^{k_{\nu+1}} \Pr(E_{\nu-1,k'}).\end{aligned} \tag{98}$$

By (9 bis)[4] we have

$$\begin{aligned}\Pr(E_{\nu-1,k'}) &\leqq \Pr\left(\max_{k'_{\nu-1}<\rho\leqq k'} |S_\rho - S_{k'_{\nu-1}}| < 8^{-1/2}\pi\frac{(s^2_{k'} - s^2_{k'_{\nu-1}})^{1/2}}{(\psi^2_{k'} + 8C)^{1/2}}\right)\\ &\geqq \frac{A_6}{A_7}\Pr\left(S^*_{k'} < 8^{-1/2}\pi\frac{s_{k'}}{(\psi^2_{k'} + 8C)^{1/2}}\right).\end{aligned}$$

Thus from (98) and the last inequality we shall have proved (95) with $A_{10}=A_6/2NA_7$. Hence it is sufficient to prove (97).

Now we have, since $\psi'^2_k \geqq \psi'^2_{k_\nu} \geqq \psi'^2_{k_{\nu+1}} - 1 \geqq \psi'^2_j - 1$,

$$\frac{s'_k}{(\psi'^2_k + 8C)^{1/2}} < \frac{s'_j}{(\psi'^2_j - 1 + 8C)^{1/2}},$$

$$\Pr(E'_{k,j}) \leqq \Pr\left(\max_{n_k \leqq i \leqq n_j} |S_i - S_{n_k}| < \frac{\pi}{8^{1/2}}(s'^2_j - s'^2_k)^{1/2} g_j\right) \tag{99}$$

where

$$g_j = \frac{2s'_j}{(s'^2_j - s'^2_k)^{1/2}(\psi'^2_j + 8C - 1)^{1/2}}. \tag{100}$$

---

(4) See footnote 3.

It is obvious that $g_j \downarrow 0$; in order to apply Theorem 2 we have to verify that

$$\frac{(s_j'^2 - s_k'^2)^{1/2}(\psi_j'^2 + 8C - 1)^{1/2}}{s_j'} \leqq A_{11}(\lg_2 (s_j'^2 - s_k'^2))^{1/2}$$

which is evident since

$$\left(\frac{s_j'^2 - s_k'^2}{\lg_2 (s_j'^2 - s_k'^2)}\right)^{1/2} \leqq A_{12} \frac{s_j'}{\psi_j'},$$

by (76). Therefore we have from (99) and (9), Theorem 2,

$$\text{(101)} \qquad \Pr (E_{k,j}') \leqq A_{13}e - g_j^{-2}.$$

We have for sufficiently large k, from (77),

$$\text{(102)} \qquad \frac{s_k'}{s_{k+1}'} \leqq 1 - \frac{a}{2\psi_k'^2}, \qquad \frac{s_k'}{s_j'} \leqq \left(1 - \frac{a}{2\psi_j'^2}\right)^{j-k}.$$

If $hx \leqq \delta$ where $\delta > 0$ is sufficiently small, then $(1-x)^h \leqq 1 - \delta' hx$ where $\delta' > 0$ is another constant. Hence if $j - k \leqq \delta\psi_j'^2$ we have from (102)

$$\frac{s_k'}{s_j'} \leqq 1 - \frac{\delta' a(j-k)}{\psi_j'^2}, \qquad 1 - \frac{s_k'^2}{s_j'^2} \geqq a' \frac{j-k}{\psi_j'^2}$$

where $a' > 0$. Then from (100)

$$g_j \leqq 2(a'(j-k))^{-1/2}.$$

Hence by (101) we have

$$\text{(103)} \qquad \Pr (E_{k,j}') \leqq A_{13} \exp (-4^{-1}a'^2(j-k)).$$

If $hx > \delta$, then $(1-x)^h < \delta'' < 1$, hence from (102), if $j - k > \delta\psi_j'^2$,

$$\frac{s_k'}{s_j'} < \delta_0 < 1, \qquad 1 - \frac{s_k'^2}{s_j'^2} \geqq 1 - \delta_0^2;$$

$$g_j \leqq 2(1 - \delta_0^2)^{-1/2}\psi_j'^{-1};$$

$$\text{(104)} \qquad \Pr (E_{k,j}') \leqq A_{13} \exp (-4^{-1}(1 - \delta_0^2)\psi_j'^2).$$

From (103) and (104),

$$\text{(105)} \quad \sum_{j=k+N}^{k_{\nu+1}} \Pr (E_{k,j}') \leqq A_{13}\left(\sum_{i=N}^{\infty} e^{-a'i/4} + (k_{\nu+1} - k_\nu) \exp\left(-\frac{1-\delta_0^2}{4}\psi_k'^2\right)\right).$$

We have by (83),

$$\psi_{k_\nu}'^3 \geqq \frac{s_{k_{\nu+1}}'}{s_{k_\nu}'} \geqq \left(1 + \frac{a}{\psi_{k_{\nu+1}-1}'^2}\right) \cdots \left(1 + \frac{a}{\psi_{k_\nu}'^2}\right) \geqq 1 + \frac{a(k_{\nu+1} - k_\nu)}{\psi_{k_{\nu+1}}'^2}.$$

Hence we have

$$k_{\nu+1} - k_\nu \leq A_{14}\psi'^5_{k_{\nu+1}}.$$

Since $\nu \neq \nu(r)$, we have $\psi'^2_{k_{\nu+1}} \leq 2\psi'^2_{k_\nu}$. Hence

$$k_{\nu+1} - k_\nu \leq 6A_{14}\psi'^5_{k_\nu};$$

$$(k_{\nu+1} - k_\nu) \exp\left(-\frac{1-\delta_0^2}{4}\psi'^2_k\right) \leq A_{15}\psi'^5_{k_\nu}e^{-A_{16}\psi'^2_{k_\nu}} = o(1).$$

Thus by choosing N sufficiently large we obtain from (105) the desired (97), if ν is sufficiently large. The proof of Lemma 9 is thus complete.

LEMMA 10. *If* $\{n_k\}$, $k=1, 2, \cdots$, *is defined by* (77), *then the series*

$$\sum_k e^{-\psi^2_{n_k}}$$

and

$$\sum_n \frac{\sigma_n^2}{s_n^2}\psi_n^2 e^{-\psi_n^2}$$

converge and diverge together.

Proof. We have

$$\frac{\sigma_n^2}{s_n^2} = 1 - \frac{s_{n-1}^2}{s_n^2}.$$

Since $x \leq -\lg(1-x) \leq 2x$ if $0<x<1$, we obtain

$$\frac{\sigma_n^2}{s_n^2} \leq -\lg\left(1-\frac{\sigma_n^2}{s_n^2}\right) = \lg\frac{s_n^2}{s_{n-1}^2},$$

$$\sum_{n_k<n\leq n_{k+1}} \frac{\sigma_n^2}{s_n^2} \leq \lg\frac{s'^2_{k+1}}{s'^2_k} \leq 2\lg\left(1+\frac{b}{\psi'^2_k}\right) \leq \frac{2b}{\psi'^2_k},$$

$$\frac{2\sigma_n^2}{s_n^2} \geq -\lg\left(1-\frac{\sigma_n^2}{s_n^2}\right) = \lg\frac{s_n^2}{s_{n-1}^2},$$

$$2\sum_{n_k<n\leq n_{k+1}} \frac{\sigma_n^2}{s_n^2} \geq \lg\frac{s'^2_{k+1}}{s'^2_k} \geq 2\lg\left(1+\frac{a}{\psi'^2_k}\right) \geq \frac{a}{\psi'^2_k}.$$

Since $\psi_n^2 e^{-\psi_n^2}\downarrow$, we have

$$\frac{a}{2}e^{-\psi'^2_{k+1}} = \psi'^2_k e^{-\psi'^2_{k+1}}\frac{a}{2\psi'^2_k} \leq \sum_{n_k<n\leq n_{k+1}} \frac{\sigma_n^2}{s_n^2}\psi_n^2 e^{-\psi_n^2} \leq \psi'^2_k e^{-\psi'^2_k}\frac{2b}{\psi'^2_k} \leq 2be^{-\psi'^2_k}.$$

Lemma 10 follows from this inequality.

Proof of Theorem 3. The ϕ_n given in (11) is monotone increasing and $\phi_n = O((\lg_2 s_n)^{1/2})$. Hence Lemma 8 and Lemma 9 are applicable. Hence

$$\Pr\,(S_n^* < 8^{-1/2}\pi s_n \phi_n^{-1} \text{ i. o.}) = \begin{cases} 0 \\ 1 \end{cases}$$

according as

$$\sum_k e^{-\psi_{n_k}^2} \begin{Bmatrix} < \\ = \end{Bmatrix} \infty.$$

By Lemma 10, the last series converges and diverges with (13), which in this case is

$$\sum_n \frac{(1 + o(1))\sigma_n^2 \lg_2 s_n}{s_n^2 \lg s_n (\lg_2 s_n)^2 \lg_3 s_n \cdots \lg_p s_n (\lg_{p+1} s_n)^{1+\delta}}.$$

Hence a well known theorem of Abel-Dini asserts that this is convergent if and only if δ is positive. Thus Theorem 3 is proved.

Proof of Theorem 4. Suppose that $\phi_n \uparrow \infty$. Define

$$\psi_n^2 = \min\,(\phi_n^2, 2 \lg_2 s_n). \tag{106}$$

If (13) is convergent, then

$$\sum_n \frac{\sigma_n^2}{s_n^2} \psi_n^2 e^{-\psi_n^2} = \sum_{\psi_n = \phi_n} + \sum_{\psi_n^2 \geqq 2 \lg_2 s_n} < \infty$$

since again by the Abel-Dini theorem we have

$$\sum_{\psi_n^2 \geqq 2 \lg_2 s_n} \leqq \sum \frac{2\sigma_n^2 \lg_2 s_n}{s_n^2 (\lg s_n)^2} < \infty.$$

By the definition (106) ψ_n satisfies (76), hence by Lemma 8,

$$\Pr\,(S_n^* < 8^{-1/2}\pi s_n \psi_n^{-1} \text{ i. o.}) = 0.$$

Since $\psi_n \leqq \phi_n$, a fortiori (12) is equal to zero.

If (13) is divergent, then since $\psi_n \leqq \phi_n$, we have

$$\sum_n \frac{\sigma_n^2}{s_n^2} \psi_n^2 e^{-\psi_n^2} = \infty.$$

Since ψ_n satisfies (76), by Lemma 9, $\Pr\,(S_n^* < 8^{-1/2}\pi s_n \psi_n^{-1} \text{ i. o.}) = 1$. By Theorem 3, we have

$$\Pr\left(S_n^* < 8^{-1/2}\pi \frac{s_n}{(2 \lg_2 s_n)^{1/2}} \text{ i. o.}\right) = 0.$$

Hence there exists a subsequence n_i such that $\psi_{n_i}^2 \leqq 2 \lg_2 s_{n_i}$ and

$$\Pr\left(S^*_{n_i} < 8^{-1/2}\pi s_{n_i}\psi^{-1}_{n_i} \text{ i. o.}\right) = 1.$$

By the definition (106), we have $\psi_{n_i} = \phi_{n_i}$. Hence (12) is equal to one. Theorem 4 is proved.

Proof of Theorem 5. By Theorem 4 it is sufficient to prove that the series

$$\sum_n \frac{\sigma_n^2}{s_n^2}\phi^2(s_n^2)e^{-\phi^2(s_n^2)} \tag{107}$$

and the integral (15) converge and diverge together.

We have, since $t^{-1}\phi^2(t)e^{-\phi^2(t)} \downarrow 0$,

$$\int_{s_k^2}^{\infty} t^{-1}\phi^2(t)e^{-\phi^2(t)}dt = \sum_{n=k+1}^{\infty}\int_{s_{n-1}^2}^{s_n^2} t^{-1}\phi^2(t)e^{-\phi^2(t)}dt$$
$$\geqq \sum_{n=k+1}^{\infty}\frac{s_n^2 - s_{n-1}^2}{s_n^2}\phi^2(s_n^2)e^{-\phi^2(s_n^2)}.$$

Hence if (107) diverges, (15) diverges too.

On the other hand, we have

$$\sum_{n=N+1}^{\infty}\frac{\sigma_n^2}{s_{n-1}^2}\phi^2(s_{n-1}^2)e^{-\phi^2(s_{n-1}^2)} \geqq \int_{s_N^2}^{\infty} t^{-1}\phi^2(t)e^{-\phi^2(t)}dt. \tag{108}$$

From (75) we have $s_n^2 = s_{n-1}^2 + \sigma_n^2 \leqq s_{n-1}^2 + O(s_n^{2-2\theta})$. Hence if n is large enough, we have

$$s_n^2 \leqq 2s_{n-1}^2. \tag{109}$$

Let n_k, $k=1, 2, \cdots$, denote the subsequence of $n=1, 2, \cdots$ for which

$$\phi^2(s_{n_k}^2) > \phi^2(s_{n_k-1}^2) + 1. \tag{110}$$

Evidently we have by (109) and (110),

$$\sum_k \frac{\sigma_{n_k}^2}{s_{n_k-1}^2}\phi^2(s_{n_k-1}^2)e^{-\phi^2(s_{n_k-1}^2)} \leqq A_{16}\sum_k \frac{\sigma_{n_k}^2}{s_{n_k}^2}e^{-\phi^2(s_{n_k}^2-1)/2} < \infty. \tag{111}$$

Hence if (15) diverges, we have, by (108) and (111),

$$\sum_{n=1, n\neq n_k}^{\infty}\frac{\sigma_n^2}{s_{n-1}^2}\phi^2(s_{n-1}^2)e^{-\phi^2(s_{n-1}^2)} = \infty. \tag{112}$$

By (110) if $n \neq n_k$, we have $\phi^2(s_n^2) \leqq \phi^2(s_{n-1}^2)+1$. From this and (112) we obtain

$$\sum \frac{\sigma_n^2}{s_n^2}\phi^2(s_n^2)e^{-\phi^2(s_n^2)} \geqq \frac{e^{-1}}{2}\sum \frac{\sigma_n^2}{s_{n-1}^2}\phi^2(s_{n-1}^2)e^{-\phi^2(s_{n-1}^2)} = \infty.$$

Theorem 5 is proved.

References

1. L. Bachelier, *Calcul des probabilités*, vol. 1, Paris, 1912.

2. H. Bergström, *On the central limit theorem in R_k, $k>1$*, Skand. Aktuarietidskr. vol. 28 (1945) pp. 106–127.

3. A. C. Berry, *The accuracy of the Gaussian approximation to the sum of independent variates*, Trans. Amer. Math. Soc. vol. 49 (1941) pp. 122–136.

4. K. L. Chung and Paul Erdös, *On the lower limit of sums of independent random variables*, Ann. of Math. vol. 48 (1947) pp. 1003–1013.

5. H. Cramér, *Random variables and probability distributions*, Cambridge Tracts in Mathematics, No. 36, Cambridge University Press, 1937.

6.——— , *Sur un nouveau théorème-limite de la théorie des probabilités*, Actualités Scientifiques, No. 736, Paris, 1938, pp. 5–23.

7. ———, *Mathematical methods of statistics*, Uppsala, 1945. (Princeton University Press, 1946.)

8. P. Erdös and M. Kac, *On certain limit theorems of the theory of probability*, Bull. Amer. Math. Soc. vol. 52 (1946) pp. 292–302.

9. C.-G. Esseen, *Fourier analysis of distribution functions*, Acta Math. vol. 77 (1945) pp. 1–125.

10. W. Feller, *Generalization of a probability limit theorem of Cramér*, Trans. Amer. Math. Soc. vol. 54 (1943) pp. 373–402.

11. ———, *The general form of the so-called law of the iterated logarithm*, Trans. Amer. Math. Soc. vol. 54 (1943) pp. 373–402.

12. ———, *The fundamental limit theorems in probability*, Bull. Amer. Math. Soc. vol. 51 (1945) pp. 800–832.

13. M. Fréchet, *Recherches théoriques modernes sur la théorie des probabilités*, vol. 1, Paris, 1937.

14. A. Khintchine, *Ein Satz der Wahrscheinlichkeitsrechnung*, Fund. Math. vol. 6 (1924) pp. 9–10.

15. A. Kolmogoroff, *Ueber die Summen durch den Zufall bestimmter unabhängiger Grössen*, Math. Ann. vol. 99 (1928) pp. 309–319.

16. ———, *Ueber das Gesetz des iterierten Logarithmus*, Math. Ann. vol. 101 (1929) pp. 126–135.

17. A. Liapounoff, *Nouvelle forme du théorème sur la limite de la probabilité*, Mémoires de l'Académie de Saint-Petersbourg (8) vol. 12 (1901).

18. J. V. Uspensky, *Introduction to mathematical probability*, New York, McGraw-Hill, 1937.

Princeton University,
Princeton, N. J.

ANNALS OF MATHEMATICS
Vol. 50, No. 2, April, 1949

ON THE ZEROS OF $\sum_1^n \pm 1$

K. L. CHUNG AND G. A. HUNT

(Received April 16, 1948)

Let $X_1, X_2, \cdots$ be independent random variables with the distribution

$$\Pr\{X_v = 1\} = 1/2 = \Pr\{X_v = -1\} \qquad v = 1, 2, \cdots .$$

We propose to establish theorems similar to the law of the iterated logarithm, but concerning the number of zeros in the sequence

$$\sum_1^n X_v \qquad n = 1, 2, \cdots$$

of partial sums rather than the magnitude of $\sum_1^n X_v$. The theorems are stated at the beginning of paragraphs 3 and 4; the estimates needed in the proofs are formulated as lemmas near the end of paragraph 1. In paragraph 5 we indicate how similar theorems concerning changes of sign can be proved.

1. The distribution of N_n and W_r

Since a partial sum of odd order cannot vanish, we set

$$S_0 \equiv 0 \qquad S_n \equiv X_1 + \cdots + X_{2n} \qquad n = 1, 2, \cdots$$

N_n denotes the number of zeros among $S_1, S_2, \cdots, S_n$, while W_r is the subscript of the r^{th} zero in the sequence $S_1, S_2, \cdots$ ad inf.—that is to say, $S_{W_r} = 0$ and $S_v = 0$ for exactly $r - 1$ values of v in the range $1 \leq v \leq W_r - 1$. The quantities N_n and W_r are in some sense equivalent, since $N_n < r$ means the same as $W_r > n$; but we shall usually find W_r the more convenient to work with.

Using the notations

$$w_k = \Pr\{W_1 = k\}$$

$$s_n = \Pr\{S_n = 0\} = \binom{2n}{n}\frac{1}{2^{2n}}$$

$$\theta(z) = \sum_1^\infty w_k z^k$$

$$T(z) = \sum_0^\infty s_n z^n = (1 - z)^{-\frac{1}{2}},$$

we have

$$s_n = \sum_{k=1}^n \Pr\{W_1 = k\} \Pr\{S_n - S_k = 0\}$$

$$= \sum_{k=1}^n w_k s_{n-k}$$

$$T(z) - 1 = \theta(z)T(z)$$

$$\theta(z) = 1 - T(z)^{-1} = 1 - (1 - z)^{1/2}.$$

Now $W_r \equiv V_1 + V_2 \cdots + V_r$, the V_k being independent and each V_k having the same distribution as W_1. So W_r has

$$\theta^r(z) = [1 - (1 - z)^{1/2}]^r$$

for its generating function. Expanding $\theta^r(z)$ in Lagrange's series, we obtain

$$\Pr\{W_r \leqq n\} = \sum_{k=1}^{n} \{\text{coefficient of } z^k \text{ in } \theta^r(z)\}$$

$$= \sum_{k=0}^{n-r} \frac{r(r + k + 1) \cdots (r + 2k - 1)}{1 \cdot 2 \cdots k} \frac{1}{2^{2k+r}},$$

where the first two terms of the last sum are to be interpreted as 2^{-r} and $r2^{-2-r}$. For the last expression, which is familiar from the problem of the gambler's ruin, we have the estimate (Uspensky [5]):

$$\Pr\{W_r \leqq n\} = \Pr\{N_n \geqq r\} = (2/\pi)^{1/2} \int_t^\infty e^{-u^2/2}\, du + \frac{\epsilon}{6n} \tag{1}$$

with $t = 2r(2n + 2/3)^{-1/2}$, $|\epsilon| < 1$, and $n \geqq 50$.

Equation (1) readily yields the lemmas below, which contain the information we need. $\psi_1, \phi_1, k_1, \cdots$ are suitable constants.

LEMMA 1. *If $\psi < \phi_1$ and $r > r_1$ then*

$$\Pr\left\{W_r > \frac{2r^2}{\psi^2}\right\} = \left(\frac{2}{\pi}\right)^{1/2} \psi(1 + \epsilon), \qquad |\epsilon| < 1/3.$$

LEMMA 1'. *If $\phi < \phi_1$ and $\phi n^{1/2} > k_1$ then*

$$\Pr\left\{N_n < \left(\frac{n}{2}\right)^{1/2} \phi\right\} = \left(\frac{2}{\pi}\right)^{1/2} \phi(1 + \epsilon), \qquad |\epsilon| < 1/3.$$

LEMMA 2. *If $\psi_2 < \psi < 3/2 \lg^{1/2} r$ then*

$$\Pr\left\{W_r < \frac{2r^2}{\psi^2}\right\} = \left(\frac{2}{\pi}\right)^{1/2} \frac{1}{\psi} e^{-\psi^2/2}(1 + \epsilon), \qquad |\epsilon| < 1/3.$$

It is interesting to note that as $r \to \infty$ the distribution of $W_r/(2r^2)$ approaches the stable distribution whose characteristic function is

$$\exp\{-\tfrac{1}{2}(1 - i \operatorname{sgn} t) |t|^{1/2}\}.$$

This may be verified by setting $z \equiv \exp\{it/(2r^2)\}$ in $\theta^r(z)$ and letting $r \to \infty$. At the same time the distribution tends to that of $1/Y^2$, where Y is a normally distributed variable with mean 0 and variance 1; and the distribution of $(n/2)^{-1/2}N_n$ tends to that of $|Y|$. The last two statements follow at once from (1). In this connection see a paper by P. Lévy [4].

2. A relation between N_n and W_r

By $\Phi(x)$ or $\Psi(y)$ we shall always mean a positive and continuous function strictly increasing to ∞, though it will be clear from the proofs that it is enough to have Φ and Ψ non-decreasing and positive.

Let us suppose for the moment that Ψ is the inverse of Φ. Then, because $N_n = r$ for $W_r \leqq r < W_{r+1}$, the statement

$$(2') \qquad N_n > \Phi(n) \text{ i.o.}$$

is equivalent to

$$r > \Phi(W_r) \text{ i.o.}$$

or to

$$(2) \qquad W_r < \Psi(r) \text{ i.o.}$$

Here i.o. stands for 'infinitely often,' that is to say, 'for infinitely many n' or 'for infinitely many r'. Similarly

$$(3') \qquad N_n < \Phi(n) \text{ i.o.}$$

is the same as

$$(3) \qquad W_r > \Psi(r-1) + 1 \text{ i.o.}$$

ϕ and ψ will always be related to Φ and Ψ by

$$\Psi(y) = \frac{2y^2}{\psi^2(y)} \qquad\qquad \Phi(x) = \left(\frac{x}{2}\right)^{1/2} \phi(x).$$

When Ψ is the inverse of Φ, we shall have

$$\phi(x) = \psi(y) \qquad x = \Psi(y) \qquad y = \Phi(x).$$

3. Lower bounds for N_n

THEOREM 1. *If* $\psi(y) \downarrow 0$ *then*

$$P_1(\psi) = \Pr\left\{W_r > \frac{2r^2}{\psi^2(r)} \text{ i.o.}\right\} = 0 \text{ or } 1$$

as

$$I_1(\psi) = \int_1^\infty \frac{\psi(y)}{y}\, dy < \infty \text{ or } = \infty.$$

THEOREM 1′. *If* $\phi(x) \downarrow 0$ *and* $\phi(x)x^{1/2} \uparrow \infty$, *then*

$$P_1'(\phi) = \Pr\left\{N_n < \left(\frac{n}{2}\right)^{1/2} \phi(n) \text{ i.o.}\right\} = 0 \text{ or } 1$$

as

$$I_1(\phi) = \int_1^\infty \frac{\phi(y)}{y}\, dy < \infty \text{ or } = \infty.$$

We make a few remarks before entering into the proof. First, it is easy to see that $I_1(\psi)$ can be replaced by the series $\sum\psi(2^k)$ and $I_1(\phi)$ by $\sum\phi(2^k)$.

Next, let us suppose that

$$y = \Phi(x) = \left(\frac{x}{2}\right)^{1/2}\phi(x) \tag{4}$$

$$x = \Psi(y) = \frac{2y^2}{\psi^2(y)}.$$

Then if ψ satisfies the hypothesis of Theorem 1, ϕ satisfies that of Theorem 1′, and conversely. Also

$$\int_1^\infty \frac{\phi(x)}{x}\,dx = -2^{3/2}\int_1^\infty \Phi(x)\,dx^{-1/2}$$

$$= 2^{3/2}\,\Phi(1) + 2^{3/2}\int_1^\infty \frac{1}{x^{1/2}}\,d\Phi(x) = 2^{3/2}\,\Phi(1) + 2\int_{\Phi(1)}^\infty \frac{\psi(y)}{y}\,dy,$$

so that the integrals $I_1(\phi)$ and $I_1(\psi)$ converge or diverge together.

Finally, keeping in mind that (3) and (3′) are equivalent when (4) holds, we see that it suffices to prove:

(a) If $I_1(\phi) < \infty$ then $P_1'(\phi) = 0$

(b) If $I_1(\psi) = \infty$ then $P_1(\psi) = 1$.

PROOF OF (a). $I_1(\phi) < \infty$ implies

$$\sum_1^\infty \phi(2^k) < \infty,$$

so that by Lemma 1′

$$\sum_1^\infty \Pr\left\{N_{2^k} < \left(\frac{2^k}{2}\right)^{1/2} 2\phi(2^k)\right\} = \sum_1^\infty \left(\frac{2}{\pi}\right)^{1/2}(1+\epsilon_k)2\phi(2^k) < \infty. \tag{5}$$

From (5) and the lemma of Borel and Cantelli we conclude that there is a random variable K, almost always finite, such that

$$N_{2^k} \geqq \left(\frac{2^k}{2}\right)^{1/2} 2\phi(2^k), \qquad \text{if } k \geqq K.$$

When n surpasses 2^K,

$$N_n \geqq N_{2^k} \geqq \left(\frac{2^k}{2}\right)^{1/2} 2\phi(2^k) \geqq \left(\frac{2^{k+2}}{2}\right)^{1/2}\phi(2^k) \geqq \left(\frac{n}{2}\right)^{1/2}\phi(n),$$

$$2^K \leqq 2^k < n \leqq 2^{k+1}.$$

So (a) is true.

Proof of (b). $I_1(\psi) = \infty$ is equivalent to

$$\sum_1^\infty \psi(2^k) = \infty.$$

Hence, according to Lemma 1,

$$\sum_k \Pr\left\{W_{2^k} - W_{2^{k-1}} > \frac{2\cdot 2^{2k}}{\psi^2(2^k)}\right\} = \sum_k \left(\frac{2}{\pi}\right)^{1/2}(1+\epsilon_k)\frac{\psi(2^k)}{4} = \infty. \tag{6}$$

The variables $W_{2^k} - W_{2^{k-1}} = V_{2^{k-1}+1} + \cdots V_{2^k}$ are independent. We conclude from (6) and the lemma of Borel that almost certainly $W_{2^k} - W_{2^{k-1}} > 2\cdot 2^{2k}\psi^{-2}(2^k)$ for infinitely many k. Then almost certainly

$$W_{2^k} \geqq W_{2^k} - W_{2^{k-1}} \geqq \frac{2\cdot 2^{2k}}{\psi^2(2^k)}$$

infinitely often. This proves (b).

Theorems 1 and 1′ are known also to Professor Feller. He proves them by applying to W_1 a theorem proved in [1].

4. Upper bounds for N_n

The proof of the next pair of theorems is much harder. The argument is essentially that used by Feller [2] in proving the general form of the law of the iterated logarithm.

Theorem 2. *If* $\psi(y) \uparrow \infty$ *and* $y/\psi(y) \uparrow \infty$ *then*

$$P_2\{\psi\} = \Pr\left\{W_r < \frac{2r^2}{\psi^2(r)}\text{ i.o.}\right\} = 0 \text{ or } 1$$

as

$$I_2(\psi) = \int_1^\infty \frac{\psi(y)}{y}e^{-\psi^2(y)/2}dy < \infty \text{ or } = \infty.$$

Theorem 2′. *If* $\phi(x) \uparrow \infty$ *then*

$$P_2'\{\phi\} = \Pr\left\{N_n > \left(\frac{n}{2}\right)^{1/2}\phi(n)\text{ i.o.}\right\} = 0 \text{ or } 1$$

as

$$I_2\{\phi\} = \int_1^\infty \frac{\phi(x)}{x}e^{-\phi^2(x)/2}dx < \infty \text{ or } = \infty.$$

The integral $I_2(\psi)$ may be replaced by

$$\sum_{n=1}^\infty \frac{\psi(n)}{n}e^{-\psi^2(n)/2},$$

and a like remark holds for $I_2(\phi)$.

The two theorems are equivalent. For we may suppose that ϕ and ψ are

related by (4), so that ϕ satisfies the hypothesis of Theorem 2′ if ψ satisfies that of Theorem 2, and conversely. Recalling that (2) is the same as (2′), we have only to show that $I_2(\psi)$ and $I_2(\phi)$ converge together. Now

$$\int_1^\infty \frac{\phi(x)}{x} e^{-\phi^2(x)/2}\,dx = \int_{\Phi(1)}^\infty \frac{\psi(y)}{\Psi(y)} e^{-\psi^2(y)/2}\,d\Psi(y)$$

$$= \int_{\Phi(1)}^\infty \psi(y)e^{-\psi^2(y)/2}\,d[\lg(2y^2) - \lg \psi^2(y)]$$

$$= 2\int_{\Phi(1)}^\infty \frac{\psi(y)}{y} e^{-\psi^2(y)/2}\,dy - 2\int_{\Phi(1)}^\infty e^{-\psi^2/2}\,d\psi$$

and the last integral is finite. Thus it is sufficient to prove Theorem 2.

PROOF FOR $I_2(\psi) < \infty$. If $\psi^2(y) \leqq \lg y$ is not true, we replace ψ by

$$\bar{\psi}(y) \equiv \min\{\psi(y), (\lg y)^{1/2}\}. \tag{7}$$

Then $I_2(\bar{\psi}) < \infty$ and it is enough to show that $P_2(\bar{\psi}) = 0$. Since $r\bar{\psi}^{-2}(r) \to \infty$ we may define a sequence r_k by the condition

$$r_k\left(1 + \frac{a}{\bar{\psi}^2(r_k)}\right) \leqq r_{k+1} < r_k\left(1 + \frac{b}{\bar{\psi}^2(r_k)}\right) \tag{8}$$

where $0 < a < b$ and r_1 is chosen sufficiently large. Feller [1] has proved that $I_2(\bar{\psi})$ converges or diverges with

$$\sum_1^\infty \frac{1}{\bar{\psi}(r_k)} e^{-\bar{\psi}^2(r_k)/2}.$$

We define

$$a_k = \min\{\bar{\psi}(r_k), (a_{k-1}^2 + 1)^{1/2}\}$$

so that $r_{k+1} < r_k(1 + ba_k^{-2})$. If $k_1 < k_2 < \cdots$ are the indices for which

$$a_k^2 = a_{k-1}^2 + 1,$$

the series

$$\sum_\nu \frac{1}{a_{k_\nu}} e^{-a_{k_\nu}^2/2}$$

converges (by the ratio test). Hence

$$\sum_k \frac{1}{a_k} e^{-a_k^2/2} = \sum_{a_k = \bar{\psi}(r_k)} + \sum_{a_k^2 = a_{k-1}^2 + 1} < \infty.$$

It is easy to see that also

$$\sum_k \frac{1}{(a_k^2 - 2b)^{\frac{1}{2}}} e^{-(a_k^2 - 2b)/2} < \infty.$$

Hence, according to Lemma 2

$$\sum_k \Pr\left\{W_{r_k} < \frac{2r_k^2}{a_k^2 - 2b}\right\} \tag{9}$$
$$= \sum \left(\frac{2}{\pi}\right)^{1/2} (1 + \epsilon_k) \frac{1}{(a_k^2 - 2b)} e^{-(a_k^2 - 2b)/2} < \infty.$$

From (9) and the lemma of Borel and Cantelli, we conclude that almost certainly

$$W_{r_k} \geq \frac{2r_k^2}{a_k^2 - 2b}$$

for all sufficiently large k. Now, if $r_k \leq r < r_{k+1}$ and $W_{r_k} \geq 2r_k^2(a^2 - 2b)^{-1}$, then

$$\begin{aligned} W_r \geq W_{r_k} &\geq \frac{2r_k^2}{a_k^2 - 2b} \\ &\geq \frac{2r^2}{\bar{\psi}^2(r)} \frac{r_k^2}{r_{k+1}^2} \frac{a_k^2}{a_k^2 - 2b} \\ &\geq \frac{2r^2}{\bar{\psi}^2(r)} \frac{1}{1 + b/a_k^2} \frac{1}{1 - 2b/a_k^2} \geq \frac{2r^2}{\bar{\psi}^2(r)} \end{aligned}$$

Thus $P_2(\bar{\psi}) = 0$.

PROOF FOR $I_2(\psi) = \infty$. First let us show that we need consider only $\psi(y)$ dominated by $\lg^{1/2} y$. We define $\bar{\psi}$ again by (7), so that $I_2(\bar{\psi}) = \infty$. Assuming the full theorem for functions less than $\lg^{1/2} y$ (we have already proved one part), we have

$$\Pr\left\{W_r < \frac{2r^2}{\bar{\psi}^2(r)} \text{ i.o.}\right\} = 1,$$

and also

$$\Pr\left\{W_r < \frac{2r^2}{\lg r} \text{ i.o.}\right\} = 0$$

since $I_2(\lg^{1/2} y) < \infty$. Letting $r_1 < r_2 < \cdots$ be the successive r's for which $\bar{\psi}(r) = \psi(r)$ we conclude that the sequence r_i is infinite and that

$$\Pr\left\{W_r < \frac{2r_i^2}{\psi^2(r_i)} \text{ i.o.}\right\} = 1.$$

The last statement is even stronger than $P_2(\psi) = 1$.

From now on we assume that $\psi^2(y) \leq \lg y$. The sequence r_k is defined as before by (8); and now

$$\sum_{k=1}^{\infty} \frac{1}{\psi(r_k)} e^{-\psi^2(r_k)/2} = \infty. \tag{10}$$

From r_k we select the subsequence r_{k_v} satisfying

$$r_{k_{v+1}} \leqq r_{k_v}\psi^3(r_{k_v}) < r_{k_{v+1}+1}\,, \tag{11}$$

k_0 being chosen so large that $k_v \to \infty$. In order to render the printer's task less irksome, we write r'_v for r_{k_v}. (8) and (11) imply

$$\psi^3(r'_v) \geqq \frac{r'_{v+1}}{r'_v} = \prod_{k=k_v}^{k_{v+1}-1} \frac{r_{k+1}}{r_k} \geqq \prod_{k_v}^{k_{v+1}-1} [1 + a/\psi^2(r_k)] \geqq a \prod_{k_v}^{k_{v+1}-1} 1/\psi^2(r_k),$$

so that for large v

$$\begin{aligned} \sum_{k=k_v}^{k_{v+1}-1} \frac{1}{\psi(r_k)}\, e^{-\psi^2(r_k)/2} &\leqq e^{-\psi^2(r'_v)/3} \sum_{k_v}^{k_{v+1}-1} \frac{1}{\psi(r_k)}\, e^{-\psi^2(r_k)/6} \\ &\leqq e^{-\psi^2(r'_v)/3} \sum_{k_v}^{k_{v+1}-1} 1/\psi^2(r_k) \\ &\leqq \frac{\psi^3(r'_v)}{a}\, e^{-\psi^2(r'_v)/3} < e^{-\psi^2(r'_v)/4}\,. \end{aligned} \tag{12}$$

Now let $v(1) < v(2) < \cdots$ be the v's for which

$$\psi^2(r_{k_{v+1}}) > \psi^2(r_{k_{v-1}}) + 1. \tag{13}$$

Clearly

$$\psi^2(r'_{v(i+2)}) \geqq \psi^2(r'_{v(i)}) + 1,$$

and hence

$$\sum_{v=v(i)} \sum_{k=k_v}^{k_{v+1}-1} \frac{1}{\psi(r_k)}\, e^{-\psi^2(r_k)/2} \leqq \sum_i e^{-\psi^2(r'_{v(i)})/4} < \infty$$

by (12) and the ratio test. Thus

$$\sum_{v\neq v(i)} \sum_{k=k_v}^{k_{v+1}-1} \frac{1}{\psi(r_k)}\, e^{-\psi^2(r_k)/2} = \infty.$$

In the following arguments we even assume

$$\sum_{\substack{v \text{ even} \\ v\neq v(i)}} \sum_{k=k_v}^{k_{v+1}-1} \frac{1}{\psi(r_k)}\, e^{-\psi^2(r_k)/2} = \infty, \tag{14}$$

or else we could reason on the corresponding divergent summation over the odd v.

Here are the definitions of a number of sets which will be used. $\{R\}$ denotes the set (in the product space of the X_v) on which the relation R is true.

$$E_r = \left\{W_r < \frac{2r^2}{\psi^2(r)}\right\}$$

$$E'_{v-1} = \left\{W_{r'_{v-1}} < \frac{2r'^2_{v-1}}{C^2}\right\}$$

$$(15)\qquad E_{v-1,r} \equiv \left\{W_r - W_{r'_{v-1}} < \frac{2r^2}{\psi^2(r)} - \frac{2r'^2_{v-1}}{C^2}\right\}, r'_v \leqq r < r'_{v+1}$$

$$F_v = \sum_{r=r'_v}^{r'_{v+1}} E_{v-1,r}$$

$$E'_{v-1,r} \equiv \left\{\frac{2r^2}{\psi^2(r)+8} - \frac{2r'^2_{v-1}}{C^2} < W_r - W_{r'_{v-1}} < \frac{2r^2}{\psi^2(r)} - \frac{2r'^2_{v-1}}{C^2}\right\}$$

$$F_{k,j} \equiv \left\{W_{r_j} - W_{r_k} < \frac{2r_j^2}{\psi^2(r_j)} - \frac{2r_k^2}{\psi^2(r_k)+8}\right\}, \qquad k < j.$$

We must prove that

$$(16)\qquad \Pr\left\{\sum_{r>r_0} E_r\right\} \equiv P_2(\psi) = 1, \qquad \text{for every } r_0.$$

We first show that (16) is implied by the statements (a) and (b) below. Then we prove (a) and (b) by means of Lemma 2. We need the folowing relations which are easy consequences of (15), the fact that $W_r = V_1 + \cdots + V_r$ is a sum of independent variables, and Lemma 2.

$$(17)\qquad \begin{array}{ll} E_{v-1,r} \text{ is independent of all } E'_{v'-1} \text{ with } v' \leqq v & \\ F_2, F_4, F_6, \cdots \quad \text{are independent} & \\ E'_{v-1}E_{v-1,r} \subset E_r & r'_v \leqq r < r'_{v+1} \\ E'_{v-1,rk}E'_{v-1,rj} \subset F_{k,j} & k < j \\ F_v \text{ is independent of all } E'_{v'-1} \text{ with } v' \leqq v & \\ \Pr\{E'_{v-1}\} > 1 - \epsilon(C) & \end{array}$$

where $\epsilon(C)$ is independent of v and tends to zero with C.

Let $r'_{v_0} > r_0$. Let $\sum^*$ denote a summation over even values of v such that $v > v_0$ and also $v \neq v(i)$; let $\sum^{**}$ denote a summation over even values of μ such that $\mu > v$ and also $\mu \neq v(i)$. Keeping (17) in mind, we see that

$$\begin{aligned} \Pr\left\{\sum_{r>r_0} E_r\right\} &\geqq \Pr\left\{\sum_{v>v_0} \sum_{r=r'_v}^{r'_{v+1}-1} E_r\right\} \\ &\geqq \Pr\left\{\sum{}^* \sum_{r=r'_v}^{r'_{v+1}-1} E_r\right\} \\ &\geqq \Pr\left\{\sum{}^* E'_{v-1} \sum_{r=r'_v}^{r'_{v+1}-1} E_{v-1,r}\right\} \\ &= \Pr\left\{\sum{}^* E'_{v-1} F_v\right\} \\ &= \Pr\left\{\sum{}^* E'_{v-1}(F_v - \sum{}^{**}F_\mu)\right\} \\ &= \sum{}^* \Pr\{E'_{v-1}\} \Pr\{F_v - \sum{}^{**}F_\mu\} \\ &\geqq [1 - \epsilon(C)] \sum{}^* \Pr\{F_v - \sum{}^{**}F_\mu\} \\ &= [1 - \epsilon(C)] \Pr\left\{\sum{}^*F_v\right\}. \end{aligned}$$

We may let $C \to 0$ to obtain

$$\Pr\{\sum_{r>r_0} E_r\} \geqq \quad \Pr\{\sum{}^* F_v\}. \tag{18}$$

Since $F_2, F_4, F_6, \cdots$ are independent sets, in order to show that the right member of (18) is one we have only to prove that

$$\sum{}^* \Pr\{F_v\} = \infty. \tag{19}$$

Once (19) is demonstrated the proof will be complete.

A point can be an element of at most N terms of the sum

$$\sum_{k=k_v}^{k_{v+1}-1} \left(E'_{v-1,r_k} - \sum_{j=k+N}^{k_{v+1}-1} E'_{v-1,r_k} E'_{v-1,r_j} \right)$$

Thus

$$\Pr\left\{ \sum_{k=k_v}^{k_{v+1}-1} E'_{v-1,r_k} - \sum_{j=k+N}^{k_{v+1}-1} E'_{v-1,r_k} E'_{v-1,r_j} \right\} \geqq \frac{1}{N} \sum_{k=k_v}^{k_{v+1}-1} \Pr\left\{ E'_{v-1,r_k} - \sum_{j=k+N}^{k_{v+1}-1} E'_{v-1,r_k} E'_{v-1,r_j} \right\}.$$

It follows that

$$\begin{aligned} \Pr\{F_v\} = \Pr\left\{ \sum_{r=r'_v}^{r'_{v+1}-1} E_{v-1,r} \right\} &\geqq \Pr\left\{ \sum_{k=k_v}^{k_{v+1}-1} E_{v-1,r_k} \right\} \geqq \Pr\left\{ \sum_{k=k_v}^{k_{v+1}-1} E'_{v-1,r_k} \right\} \\ &= \Pr\left\{ \sum_{k=k_v}^{k_{v+1}-1} E'_{v-1,r_k} - \sum_{j=k+N}^{k_{v+1}-1} E'_{v-1,r_k} E'_{v-1,r_j} \right\} \\ &\geqq \frac{1}{N} \sum_{k=k_v}^{k_{v+1}-1} \Pr\left\{ E'_{v-1,r_k} - \sum_{j=k+N}^{k_{v+1}-1} E'_{v-1,r_k} E'_{v-1,r_j} \right\} \\ &\geqq \frac{1}{N} \sum_{k=k_v}^{k_{v+1}-1} \Pr\left\{ E'_{v-1,r_k} - \sum_{j=k+N}^{k_{v+1}-1} F_{k,j} \right\} \\ &\geqq \frac{1}{N} \sum_{k=k_v}^{k_{v+1}-1} \Pr\{E'_{v-1,r_k}\} \left[1 - \sum_{j=k+N}^{k_{v+1}-1} \Pr\{F_{k,j}\} \right]. \end{aligned} \tag{20}$$

We are going to show:

(a) There is a constant $A > 0$ such that for all large v ($v \neq v(i)$) and $k_v \leqq k < k_{v+1}$

$$\Pr\{E'_{v-1,r_k}\} > A/\psi(r_k) e^{-\psi^2(r_k)/2}.$$

(b) N can be chosen once for all so that for all large v ($v \neq v(i)$) and $k_v \leqq k < k_{v+1}$

$$\sum_{j=k+N}^{k_{v+1}-1} \Pr\{F_{k,j}\} < \tfrac{1}{2}.$$

It is clear that (a), (b), (20) and (14) together imply (19) and thus the theorem.

The proof (a) follows quickly from Lemma 2. We write

$$\begin{aligned}\Pr\{E'_{v-1,r}\} &= \Pr\left\{\frac{2r^2}{\psi^2(r)+8} - \frac{2r'^2_{v-1}}{C^2} < W_r - W_{r'_{v-1}} < \frac{2r^2}{\psi^2(r)} - \frac{2r'^2_{v-1}}{C^2}\right\} \\ &= \Pr\left\{W_r - W_{r'_{v-1}} < \frac{2(r-r'_{v-1})^2}{\psi_1^2}\right\} - \Pr\left\{W_r - W_{r'_{v-1}} \leqq \frac{2(r-r'_{v-1})^2}{\psi_2^2}\right\}\end{aligned} \tag{21}$$

with

$$\frac{1}{\psi_1^2} = \frac{r^2}{\psi^2(r)(r-r'_{v-1})^2} - \frac{r'^2_{v-1}}{C^2(r-r'_{v-1})^2} \geqq \frac{1}{\psi^2(r)} - \frac{1}{C^2(r/r'_{v-1}-1)^2}$$

$$\frac{1}{\psi_2^2} = \frac{r^2}{(\psi^2(r)+8)(r-r'_{v-1})^2} - \frac{r'^2_{v-1}}{C^2(r-r'_{v-1})^2} \leqq \frac{1}{(\psi^2(r)+8)(1-r'_{v-1}/r)^2}.$$

For large v the ratio

$$\frac{\psi(r)}{\psi(r'_{v-1})} \qquad r'_v \leqq r < r'_{v+1} \tag{22}$$

is nearly one since v does not satisfy (13). Thus (8) and (11) imply

$$r/r'_{v-1} \geqq \frac{r'_v}{r'_{v-1}} = \frac{r'_v}{r_{k_v+1}}\frac{r_{k_v+1}}{r'_{v-1}} \geqq \frac{1}{1+b/\psi^2(r'_v)}\psi^3(r_{v-1}) \geqq \tfrac{1}{2}\psi^3(r_{v-1}) \geqq \tfrac{1}{3}\psi^3(r)$$

$$\frac{1}{C^2(r/r'_{v-1}-1)^2} \leqq \frac{1}{C^2(\frac{1}{3}\psi^3(r)-1)^2} < \frac{1}{4\psi^4(r)}$$

$$\psi_1 < \left[\frac{1}{\psi^2(r)} - \frac{1}{4\psi^4(r)}\right]^{-\frac{1}{2}} < (\psi^2(r)+1)^{\frac{1}{2}} \tag{23}$$

also

$$\begin{aligned}\psi_2 &> (1-r'_{v-1}/r)(\psi^2(r)+8)^{1/2} \\ &> (\psi^2(r)+2)^{1/2}\end{aligned} \tag{24}$$

$$\psi_1 < \psi_2 < 2\psi(r'_{v-1}) < 2\,\mathrm{lg}^{1/2} r'_{v-1} < [\mathrm{lg}\,(r-r'_{v-1})]^{1/2}.$$

The last three inequalities justify the use of Lemma 2 in estimating the last member of (21). The result is, according to (23) and (24),

$$\Pr\{E'_{v-1,r}\} \geqq \left(\frac{2}{\pi}\right)^{\frac{1}{2}}(1+\epsilon_1)\frac{1}{(\psi^2(r)+1)^{\frac{1}{2}}}e^{-(\psi^2(r)+1)/2}$$
$$- \left(\frac{2}{\pi}\right)^{\frac{1}{2}}(1+\epsilon_2)\frac{1}{(\psi^2(r)+2)^{\frac{1}{2}}}e^{-(\psi^2(r)+2)/2} \geqq A\frac{1}{\psi(r)}e^{-\psi^2(r)/2}$$

with $A > 0$. Thus (a) is proved.

In proving (b) we again consider only those v large enough to make the ratio (22) close to unity. For $k_v \leqq k < j \leqq k_{v+1}$

$$\Pr\{F_{k,j}\} = \Pr\left\{W_{r_j} - W_{r_k} < \frac{2r_j^2}{\psi^2(r_j)} - \frac{2r_k^2}{\psi^2(r_k)+8}\right\}$$

$$(25)\qquad \begin{aligned} &\leq \Pr\left\{W_{r_j} - W_{r_k} < \frac{2r_j^2}{\psi^2(r_j)} - \frac{2r_k^2}{\psi^2(r_k)}\left(1 - \frac{4}{\psi^2(r_k)}\right)\right\} \\ &= \Pr\left\{W_{r_j} - W_{r_k} < \frac{2(r_j^2 - r_k^2)}{\psi^2(r_k)} + \frac{8r_k^2}{\psi^4(r_k)}\right\} \\ &= \Pr\left\{W_{r_j} - W_{r_k} < \frac{2(r_j - r_k)^2}{\psi^2}\right\} \end{aligned}$$

with

$$(26)\qquad \begin{aligned} \frac{1}{\psi^2} &= \frac{r_j + r_k}{\psi^2(r_k)(r_j - r_k)} + \frac{4r_k^2}{\psi^2(r_k)(r_j - r_k)^2} \\ &\leq \frac{2}{\psi^2(r_k)(1 - r_k/r_j)} + \frac{4}{\psi^4(r_k)(r_j/r_k - 1)^2} \end{aligned}$$

where, as elsewhere, a symbol printed as r_k/r_j is always meant to be bracketed e.g. $1 - r_k/r_j$ reads $1 - (r_k/r_j)$.
First we must prove that ψ satisfies the hypotheses of Lemma 2. From (8)

$$(27)\qquad \begin{aligned} \frac{r_j}{r_k} = \prod_{h=k}^{j-1} \frac{r_{h+1}}{r_h} &\geq \prod_{k}^{j-1} [1 + a/\psi^2(r_h)] \\ &\geq [1 + a/\psi^2(r'_{\nu+1})]^{j-k} \\ &\geq 1 + \tfrac{1}{2}\, a(j-k)/\psi^2(r_k). \end{aligned}$$

So

$$(28)\qquad \frac{4}{\psi^4(r_k)(r_j/r_k - 1)^2} \leq \frac{4}{a^2(j-k)^2} \leq \frac{4}{a^2N^2}.$$

If $r_j/r_k > 3/2$, then

$$(29)\qquad \frac{2}{\psi^2(r_k)(1 - r_k/r_j)} < 3/\psi^2(r_k).$$

On the other hand, if $r_j/r_k < 3/2$ then

$$(30)\qquad 1 - \frac{r_k}{r_j} \geq \frac{2}{3}\frac{a(j-k)}{\psi^2(r_k)}$$

and

$$\frac{2}{\psi^2(r_k)(1 - r_k/r_j)} \leq \frac{3}{aN}.$$

(26)–(30) show that ψ can be made large by choosing N large and then considering only large v. Also (27) implies

$$\begin{aligned}3/2 \lg (r_j - r_k) &= 3/2 \lg r_k + 3/2 \lg (r_j/r_k - 1)\\ &\geqq 3/2\, \psi^2(r_k) + 3/2 \lg (a(j-k)/(2\psi^2(r_k)))\\ &\geqq 3/2\, [\psi^2(r_k) - \lg \psi^2(r_k)] + 3/2 \lg (a(j-k)/2)\\ &\geqq \psi^2(r_k) \geqq \psi^2.\end{aligned}$$

Thus Lemma 2 may be used to estimate the last term in (25).

Fix some small $\epsilon > 0$. We consider two cases. If $a(j-k)\psi^{-2}(r_k) < \epsilon$, then (27) implies

$$\frac{r_k}{r_j} \leqq 1 - \frac{1}{4}\frac{a(j-k)}{\psi^2(r_k)}$$

and hence

$$\frac{1}{\psi^2} \leqq \frac{2}{a(j-k)} + \frac{4}{a^2(j-k)^2} \leqq \frac{A_1}{j-k}$$

where A_1 is a positive constant. Thus

$$\text{(31)} \qquad \Pr\{F_{k,j}\} < \frac{A_2}{(j-k)^{\frac{1}{2}}} e^{-A_3(j-k)} \quad \text{for} \quad \frac{a(j-k)}{\psi^2(r_k)} < \epsilon$$

with A_2 and A_3 positive constants.

If $a(j-k)\psi^{-2}(r_k) \geqq \epsilon$, then

$$r_k/r_j < \prod_{h=k}^{j-1} (1 - 2^{-1} a\psi^{-2}(r_k)) < 1 - \epsilon_1$$

$$r_j/r_k > 1 + \epsilon_2$$

where ϵ_1 and ϵ_2 are positive constants. Thus

$$\frac{1}{\psi^2} \leqq \frac{2}{\epsilon_1 \psi^2(r_k)} + \frac{4}{\epsilon_2^2 \psi^4(r_k)} \leqq \frac{A_4}{\psi^2(r_k)}$$

with A_4 a constant. Then

$$\text{(32)} \qquad \begin{aligned}\Pr\{F_{k,j}\} &< \frac{A_5}{\psi(r_k)} e^{-A_6\psi^2(r_k)}\\ &< A_7 e^{-A_8\psi^2(r_k)}.\end{aligned}$$

Together (31) and (32) yield

$$\begin{aligned}\sum_{j=k+N}^{k_{\nu+1}-1} \Pr\{F_{k,j}\} &= \sum_{(a(j-k)/\psi^2(r_k))<\epsilon} + \sum_{(a(j-k)/\psi^2(r_k))\geqq\epsilon}\\ &\leqq \sum_{j=k+N}^{\infty} \frac{A_2}{(j-k)^{\frac{1}{2}}} e^{-A_3(j-k)} + \sum_{j=k_\nu}^{k_{\nu+1}-1} A_7 e^{-A_8\psi^2(r_k)}\\ &\leqq A_9 e^{-A_3 N} + A_7(k_{\nu+1} - k_\nu) e^{-A_8\psi^2(r_k)}.\end{aligned}$$

In order to bound the last term we recall the definition (11) of k_v. This gives

$$\psi^3(r_v') \geq \frac{r_{v+1}'}{r_v'} \geq \prod_{k=k_v}^{k_{v+1}-1} [1 + a/\psi^2(r_k)]$$
$$\geq \frac{a(k_{v+1} - k_v)}{\psi^2(r_{v+1}')}.$$

Since v is not a $v(i)$

$$k_{v+1} - k_v \leq \psi^6(r_v')$$

for large v. Thus

$$\sum_{j=k+N}^{k_{v+1}-1} P\{F_{k,j}\} \leq A_9 e^{-A_8 N} + A_7 \psi^6(r_v') e^{-A_8 \psi^2(r_v')} < \tfrac{1}{2}$$

if only N is chosen large and v is large. The proof is now complete.

5. Changes of sign

Professor Feller, in his survey [3] of the limit theorems of probability, mentioned the problem of determining the asymptotic number of 'cycles' or of changes of sign in the sequence

$$\sum_1^n Y_v \qquad n = 1, 2, \cdots$$

where the Y_v are quite general random variables. The problem is apparently difficult even when the Y_v take on only integral values. As this is the problem, however, which led to the present paper, we feel in duty bound to indicate how one can derive the properties of the number of changes of sign for the particular random variable X_v.

We keep the notation of Chapter 1. Let

$$\epsilon_r = X_{2W_r+1} X_{2W_r}$$

and define $2U$ to be the index of the first change of sign in the sequence

$$\sum_1^n X_v \qquad n = 1, 2, \cdots. \tag{33}$$

Then clearly U is the first W_r for which $\epsilon_r = +1$. So

$$\begin{aligned}\Pr\{U = k\} &= \Pr\{W_1 = k \;\&\; \epsilon_1 = +1\}\\ &\quad + \Pr\{W_2 = k \;\&\; \epsilon_1 = -1 \;\&\; \epsilon_2 = +1\}\\ &\quad + \cdots\cdot\\ &= \tfrac{1}{2}\Pr\{W_1 = k\} + \frac{1}{2^2}\Pr\{W_2 = k\} + \cdots\\ &= \tfrac{1}{2}\{\text{coeff of } z^k \text{ in } \theta\} + \frac{1}{2^2}\{\text{coeff of } z^k \text{ in } \theta^2\} + \cdots\end{aligned}$$

$$= \text{coefficient of } z^k \text{ in } \sum_{l=1}^{\infty} \theta(z)/2^l$$
$$= \text{coefficient of } z^k \text{ in } \theta(z)/(2 - \theta(z))$$
$$= \text{coefficient of } z^k \text{ in } \theta^2(z)/z .$$

Thus

$$\Pr\{U = k\} = \Pr\{W_2 = k + 1\}.$$

The index U_r of the r^{th} change of sign is the sum of r independent variables each of which has the same dstribution as U. If N_n^* denotes the number of changes of sign in the sequence (33) with indices not greater than $2n$, then clearly

$$\begin{aligned} \Pr\{N_n^* \geqq r\} &= \Pr\{U_r \leqq n\} \\ &= \Pr\{W_{2r} \leqq n + r\} \\ &= \Pr\{N_{n+r} \geqq 2r\}. \end{aligned} \tag{34}$$

(1) and (34) together give the approximation to the distribution of U_r or of N_n^*. The analogues of Lemma 1 and 2 are

LEMMA 1″. *If $\psi < \psi_1''$ and $r > r_1''$ then*

$$\Pr\{U_r > 8r^2/\psi^2\} = K_1(\psi)\psi(r)$$

where $$(2/\pi)^{1/2}2/3 < K_1(\psi) < 4/3\ (2/\pi)^{1/2}.$$

LEMMA 2″. *If $\psi > \psi_2''$ and $\psi^2 < 3/2 \lg r$ then*

$$\Pr\{U_r < 8r^2/\psi^2\} = K_2(\psi)/\psi e^{-\psi^2/2}$$

where $$(2/\pi)^{1/2}2/3 < K_2(\psi) < 4/3(2/\pi)^{1/2}.$$

These lemmas have the same form as before. Thus Theorem 1 holds if W_r is replaced by U_r. Theorem 2 becomes Theorem 2″: If $\psi(y) \uparrow \infty$ and $y/\psi(y) \uparrow \infty$ then

$$\Pr\left\{U_r < \frac{8r^2}{\psi^2(r)} \text{ i.o.}\right\} = 0 \text{ or } 1$$

as

$$\int_1^{\infty} \frac{\psi(y)}{y} e^{-(\psi^2(y)/2)}\, dy < \infty \text{ or } = \infty.$$

Since N_n^* is related to U_r just as N_n was to W_r, Theorem 1′ holds when N_n^* replaces N_n and Theorem 2′ becomes Theorem 2‴: If $\phi(n) \uparrow \infty$ then

$$\Pr\{N_n^* < (n/8)^{1/2}\phi(n)\} = 0 \text{ or } 1$$

as

$$\int_1^{\infty} \frac{\phi(x)}{x} e^{-\frac{1}{2}\phi^2(x)}\, dx < \infty \text{ or } = \infty.$$

PRINCETON UNIVERSITY

[1] W. Feller, *A Limit Theorem for Random Variables with Infinite Moments*. Amer. J. Math. 68 (1946) pp. 257-262.
[2] W. Feller, *The General Form of the so-called Law of the Iterated Logarithm*, Trans. Amer. Math. Soc. 54 (1943) pp. 373–402.
[3] W. Feller, *The Fundamental Limit Theorems in Probability*. Bull. Amer. Math. Soc. 51 (1945) 800-832.
[4] P. Lévy, *Sur certains processus stochastiques homogenes*. Compositio Math. 7 (1939) pp. 283-339.
[5] J. V. Uspensky, Introduction to Mathematical Probability, New York, McGraw-Hill, 1937.

Annals of Mathematics
Vol. 51, No. 3, May, 1950

FLUCTUATIONS OF SUMS OF INDEPENDENT RANDOM VARIABLES

By Kai Lai Chung[1]

(Received March 16, 1949)

1. One aspect of the theory of addition of independent random variables is the frequency with which the partial sums change sign. Investigations of this nature were originated by Paul Lévy, in a paper [1] which contains a wealth of ideas. This problem as such was mentioned by Feller in his 1945 address [2]. In the case where the partial sums can actually vanish the problem falls under the head of 'recurrent events', a general theory of which was recently developed in a paper by Feller [3]. A very special case had been studied in detail by Hunt and myself [4].

Generalizing the problem in a natural way we shall consider the number of times T_n with which the sequence of reduced partial sums $S_k - E(S_k)$, $k = 1, 2, \cdots, n$ crosses a given value c. We shall establish the limiting distribution of T_n in the case where the random variables have a common distribution with a finite third absolute moment.[2] In all cases the limiting distribution is that of the 'positive normal', but the proper normalizing factor depends on the case. A distribution function is said to be of the 'lattice type' if and only if it is a step function with all its discontinuities located at the multiples of a certain number. Without loss of generality we may suppose that this number is an integer. The minimum distance between two discontinuities is called the 'span'. A complication arises when the given value c is a possible value of $S_k - E(S_k)$. In the case of a non-lattice distribution this makes no difference but in the case of a lattice distribution it does. In fact, the meaning of the phrase '$S_k - E(S_k)$ crosses the value c' becomes ambiguous and the definition used in Theorem 2 must be regarded as merely convenient. However, other possible definitions are subject to the same treatment.

We leave open the problem of obtaining strong limit theorems for T_n; presumably they must be of the same form as those in [4].

We denote the k^{th} moment of $F(x)$ by α_k, the k^{th} absolute moment by β_k. For brevity we write α for α_1, and σ^2 for $\alpha_2 - \alpha_1^2$.

2. Let $X_1, \cdots, X_n$ be independent random variables with the common distribution function $F(x)$. We shall assume first that $F(x)$ is not of the lattice type, and that $\beta_3 < \infty$; thus we may suppose without loss of generality that

$$\alpha = 0, \alpha_2 = 1.$$

As usual we write $S_n = \sum_{k=1}^{n} X_k$. Let y_0 be any real number. We say that

[1] Research done in connection with an ONR project.

[2] Professor Kac informed me that he and Spiegel had a different method which is applicable to an absolutely continuous distribution function with further restrictions on the analytical behavior of its characteristic function.

S_k crosses the value $-y_0$ from above if $S_k > -y_0$ and $S_{k+1} < -y_0$. Define a sequence of random variables Y_n as follows:

$$Y_k = \begin{cases} 1 & \text{if} \quad S_k > -y_0 \quad \text{and} \quad S_{k+1} < -y_0 \\ 0 & \qquad\qquad \text{otherwise.} \end{cases}$$

We write $T_n = \sum_{k=1}^{n} Y_k$. Then T_k is a random variable which is the number of times the sequence S_k, $k = 1, \cdots, n$ crosses the value $-y_0$ from above, except for a possible ambiguity at S_n.

We shall denote by $F_k(x)$ the distribution function of S_k, thus $F_k(x)$ is the k^{th} iterated convolution of $F(x)$ with itself.

Let us calculate the mean of T_n.

$$E(T_n) = \sum_{k=1}^{n} E(Y_k) = \sum_{k=1}^{n} P(S_k > -y_0, S_{k+1} < -y_0).$$

Interchanging the sum and the integral and breaking the integral into two parts, we have

$$\int_{-\infty}^{0-} \sum_{k=1}^{n} \{F_k(-y_0 - x) - F_k(-y_0)\}\, dF(x) = \int_{-n^{1/6+\epsilon}}^{0-} + \int_{-\infty}^{-n^{1/6+\epsilon}}.$$

Now we use a theorem of Esseen on asymptotic expansion (Esseen [5] Theorem 2 p. 49). We have,

$$F_k(x) = \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^{xk^{-1/2}} e^{-y^2/2}\, dy + \frac{\alpha_3}{6(2\pi k)^{1/2}} \left(1 - \frac{x^2}{k}\right) e^{-x^2/2k} + o(k^{-1/2}).$$

It follows that

$$(1) \quad F_k(-y_0 - x) - F_k(-y_0) = \frac{1}{(2\pi)^{1/2}} \int_{-y_0 k^{-1/2}}^{(-y_0 - x)k^{-1/2}} e^{-y^2/2}\, dy + o(k^{-1/2})$$

where the o-term does not depend on x.

Let $\epsilon > 0$ be a sufficiently small number. For $n^{1/2-\epsilon} \leqq k \leqq n$ and $-n^{1/6+\epsilon} \leqq x < 0$, we have

$$|x|\, k^{-1/2} \leqq \frac{n^{1/6+\epsilon}}{n^{1/4-\epsilon}} = o(1).$$

Hence the estimate (1) gives, if $x < 0$

$$(1 \text{ bis}) \qquad F_k(-y_0 - x) - F_k(-y_0) = \frac{|x|}{(2\pi)^{1/2}} + o(k^{1/2}).$$

Thus

$$\int_{-n^{1/6+\epsilon}}^{0-} \sum_{k=n^{1/2-\epsilon}} \{F_k(-y_0 - x) - F_k(-y_0)\}\, dF)x)$$

$$= \int_{-n^{1/6+\epsilon}}^{0-} \left\{ \frac{|x|}{(2\pi)^{1/2}} \sum_{n^{1/2-\epsilon} \leqq k \leqq n} k^{-1/2} + o\left(\sum_{k=1}^{n} k^{-1/2}\right)\right\} dF(x)$$

$$= \left(\frac{2n}{\pi}\right)^{1/2} \int_{-n^{1/6+\epsilon}}^{0-} |x| \, dF(x)$$

$$= \left(\frac{2n}{\pi}\right)^{1/2} \frac{\beta_1}{2} + o(n^{1/2}).$$

On the other hand, it is obvious that, the same integral with the sum over $1 \leqq k < n^{1/2-\epsilon}$ is $o(n^{1/2})$. Moreover, by Tchebychev inequality,

$$\int_{-\infty}^{-n^{1/6+\epsilon}} \sum_{k=1}^{n} \{F_k(-y_0 - x) - F_k(-y_0)\} \, dF(x)$$

$$\leqq n \int_{-\infty}^{-n^{1/6+\epsilon}} dF(x) \leqq n \frac{\beta_3}{n^{1/2+3\epsilon}} = o(n^{1/2}).$$

Hence we obtain

$$E(T_n) = \frac{\beta_1}{2}\left(\frac{2n}{\pi}\right)^{1/2} + o(n^{1/2}). \tag{2}$$

Now we are going to calculate the higher moments of T_n. By the multinomial theorem we have

$$\begin{aligned} E(T_n^m) &= E\left(\left(\sum_{k=1}^{n} Y_n\right)^m\right) \\ &= \sum_{l=1}^{m} \sum_{m_1+\cdots+m_l=m} \frac{m!}{m_1! \cdots m_l!} \sum_{0<n_1<\cdots<n_l<n} E(Y_{n_1}^{m_1} \cdots Y_{n_l}^{m_l}). \end{aligned} \tag{3}$$

We have

$$\begin{aligned} \sum_{0<n_1<\cdots<n_l<n} E(Y_{n_1}^{m_1} \cdots Y_{n_l}^{m_l}) &= \sum_{0<n_1<\cdots<n_l<n} E(Y_{n_1} \ldots Y_{n_l}) \\ &= \sum P(S_{n_1} > -y_0, S_{n_1+1} < -y_0, S_{n_2} > -y_0, S_{n_2+1} < -y_0, \cdots, \\ &\qquad S_{n_l} > -y_0, S_{n_l+1} < -y_0) \end{aligned} \tag{4}$$

where the last sum runs over all possible combinations of $n_1, \cdots, n_l$ between 0 and n such that $n_i \geqq n_{i-1} + 2$ for all i, since $n_i = n_{i-1} + 1$ is impossible.

Using the independence and the equi-distribution of the X's, the last written probability can be expressed as

$$\int_{x_1=0+}^{\infty} dF_{n_1-n_0-1}(-y_0 + x_1) \int_{y_1=-\infty}^{0-} dF(-x_1 + y_1) \int_{x_2=0+}^{\infty} dF_{n_2-n_1-1}(-y_1 + x_1)$$

$$\cdot \int_{y_2=-\infty}^{0-} dF(-x_2 + y_2) \cdots \int_{x_l=0+}^{\infty} dF_{n_l-n_{l-1}-1}(-y_{l-1} + x_l) \int_{y_l=-\infty}^{0-} dF(-x_l + y_l),$$

where $n_0 = -1$.

For a fixed integer $n_0 > 0$ and a fixed number y_0, let us write

$$W(n_0, n, y_0, l) = \sum \prod_{i=1}^{l} \int_{x_i=0+}^{\infty} dF_{n_i-n_{i-1}-1}(-y_{i-1} + x_i) \int_{y_i=-\infty}^{0-} dF(-x_i + y_i)$$

where the sum runs over all possible combinations $n_1, \cdots, n_l$ between n_0 and n such that $n_i \geqq n_{i-1} + 2$. By (2) or something completely similar, we have for every y_0,

$$W(n_0, n, y_0, 1) \sim \beta_1 \left(\frac{n - n_0}{2\pi}\right)^{1/2} \tag{5}$$

We shall use induction on l. For this purpose we assume that there is a constant c_l such that

$$W(n_0, n, y_0, l) \sim c_l(n - n_0)^{1/2} \tag{6}$$

Then we have

$$\begin{aligned} &W(n_0, n, y_0, l + 1) \\ &= \sum_{n_0 < n_1 < n} \int_{x_1 = 0+}^{\infty} dF_{n_1 - n_0 - 1}(-y_0 + x_1) \int_{y_1 = -\infty}^{0-} dF(-x_1 + y_1) \cdot W(n_1, n, y_0, l) \\ &= \sum_{n_0 < n_1 < n} \int_{x_1 = -\infty}^{0-} \{F_{n_1 - n_0 - 1}(-y_0 + x_1) - F_{n_1 - n_0 - 1}(-y_0)\} \, dF(x_1) \cdot W(n_1, n, y_0, l) \end{aligned}$$

on integrating by parts and changing the variable.

We are going to break up the integral and also the sum as before. Writing $n_1 - n_0 - 1 = k$, we have

$$\begin{aligned} &\int_{x = -(n-n_0)^{1/6+\epsilon}}^{0-} \sum_{0 \leqq k \leqq (n-n_0)^{1/2-\epsilon}} \{F_k(-y_0 + x) - F_k(-y_0)\} \, dF(x) \\ &\cdot W(k + n_0 + 1, n, y_0, l) = 0((n - n_0)^{1/2-\epsilon}(n - n_0)^{l/2}) = o((n - n_0)^{(l+1)/2}) \end{aligned}$$

by the assumption (6).

Next, for $-(n - n_0)^{1/6+\epsilon} < x < 0$, $(n - n_0)^{1/2-\epsilon} - 1 < k < n - n_0 - 1$, we have

$$| x | k^{-1/2} = o(1).$$

Hence using (1 bis) and (6) we have

$$\begin{aligned} &\sum_{(n-n_0)^{1/2-\epsilon} < k+1 < n-n_0} \{F_k(-y_0 + x) - F_k(-y)\} \cdot W(k + n_0 + 1, n, y_0, l) \\ &\qquad \sim \sum_{(n-n_0)^{1/2-\epsilon} < k+1 < n-n_0} \frac{| x |}{(2\pi k)^{1/2}} c_l(n - n_0 - 1 - k)^{1/2} \\ &\qquad \sim \frac{c_l | x |}{(2\pi)^{1/2}} \int_{(n-n_0)^{1/2-\epsilon}}^{n-n_0} y^{-1/2}(n - n_0 - 1 - y)^{1/2} \, dy. \end{aligned}$$

An obvious change of variable and the formula for beta function gives

$$\sim \frac{c_l | x |}{(2\pi)^{1/2}} \frac{\Gamma\left(\frac{l+2}{2}\right)\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{l+3}{2}\right)} (n - n_0)^{(l+1)/2}.$$

Thus we obtain

$$\int_{x=-(n-n_0)^{1/6+\epsilon}}^{0-} \sum_{(n-n_0)^{1/2-\epsilon}<k+1<n-n_0} \{F_k(-y_0+x) - F_k(y_0)\}\cdot$$

$$W(k+n_0+1, n, y_0, l)\, dF(x)$$

$$\sim \frac{\beta_1 c_l \Gamma\left(\dfrac{l+2}{2}\right)}{2^{3/2}\Gamma\left(\dfrac{l+3}{2}\right)} (n-n_0)^{(l+1)/2}.$$

Finally we have

$$\int_{x=-\infty}^{-(n-n_0)^{1/6+\epsilon}} \sum_{0<k<n-n_0-1} \{F_k(-y_0+x) - F_k(-y_0)\} W(k+n_0+1, n, y_0, l)\, dF(x)$$

$$= O\left((n-n_0)(n-n_0)^{l/2} \frac{\beta_3}{(n-n_0)^{1/2+3\epsilon}}\right) = o((n-n_0)^{(l+1)/2}).$$

Altogether we have

$$W(n_0, n, y_0, l+1) \sim \frac{\beta_1 c_l \Gamma\left(\dfrac{l+2}{2}\right)}{2^{3/2}\Gamma\left(\dfrac{l+3}{2}\right)} (n-n_0)^{(l+1)/2}.$$

Therefore (6) holds for $l+1$ with

$$c_{l+1} = \frac{\beta_1 \Gamma\left(\dfrac{l+2}{2}\right)}{2^{3/2}\Gamma\left(\dfrac{l+3}{2}\right)} c_l. \tag{7}$$

From (2) or (5) we see that

$$c_1 = \frac{\beta_1}{(2\pi)^{1/2}}.$$

This and the recurrence relation (7) determines c_l:

$$c_l = \left(\frac{\beta_1}{2^{3/2}}\right)^l \frac{1}{\Gamma\left(\dfrac{l}{2}+1\right)}. \tag{8}$$

To summarize, we have

$$\sum_{0<n_1<\cdots<n_l<n} E(Y_{n_1}^{m_1} \cdots Y_{n_l}^{m_l}) = W(-1, n, y_0, l) \sim \left(\frac{\beta_1}{2^{3/2}}\right)^l \frac{1}{\Gamma\left(\dfrac{l}{2}+1\right)} n^{l/2}.$$

Substituting into (3) we see that the term $l = m$ predominates all the others, and thus

$$E(T_n^m) \sim \left(\frac{\beta_1}{2^{3/2}}\right)^m \frac{\Gamma(m+1)}{\Gamma\left(\dfrac{m}{2}+1\right)} n^{m/2}$$

$$= \frac{2^{m/2}\Gamma\left(\dfrac{m+1}{2}\right)}{\Gamma(\frac{1}{2})} \left(\frac{\beta_1}{2}\right)^m n^{m/2}$$

by Legendre's duplication formula for the gamma function.

A normalizing factor immediately suggests itself and we obtain, for every integer $m > 0$,

$$\lim_{n\to\infty} F\left(\left(\frac{T_n}{\frac{\beta_1}{2}\, n^{1/2}}\right)^m\right) = \frac{2^{m/2}\,\Gamma\left(\dfrac{m+1}{2}\right)}{\Gamma(\frac{1}{2})} = \left(\frac{2}{\pi}\right)^{1/2} \int_0^\infty y^m e^{-y^2/2}\, dy.$$

Hence, by the continuity theorem of the moments problem (Pólya [6]), we conclude that

$$\lim_{n\to\infty} P\left(T_n \leqq \frac{\beta_1}{2}\, n^{1/2}\, x\right) = \left(\frac{2}{\pi}\right)^{1/2} \int_0^x e^{-y^2/2}\, dy. \tag{9}$$

We state this result as follows:

THEOREM 1. *Suppose $F(x)$ is not of the lattice type and*

$$\alpha = 0, \qquad \sigma^2 = 1, \qquad \beta_3 < \infty.$$

Let c be a given real number and let T_n be the number of k's not exceeding n for which

$$S_k > c, \qquad S_{k+1} < c \tag{10}$$

then

$$\lim_{n\to\infty} P\left(T_n \leqq \frac{\beta_1}{2}\, n^{1/2} x\right) = \left(\frac{2}{\pi}\right)^{1/2} \int_0^x e^{-y^2/2}\, dy.$$

Needless to say that *the number of k's not exceeding n for which*

$$S_k < c, \qquad S_{k+1} > c$$

is $T_n \pm 1$. Moreover, it turned out from the proof given above that *if one or both of the strict inequalities be replaced by $\geqq$ or $\leqq$ respectively the same result holds.*

3. We now come to the case of a lattice distribution $F(x)$. By choosing the origin and the unit of length we may suppose that the span is 1 and that

$$0 \leqq \alpha < 1, \qquad \sigma^2 > 0.$$

Two cases present themselves according as α is irrational or rational. If α is irrational then the sequence $k\alpha + c$, $k = 1, 2, \cdots$ contains at most one integer, say for $k = k_0$, thus S_k cannot take the value $k\alpha + c$ except for $k = k_0$. Then we define for $k \neq k_0 - 1$ or k_0,

$$Y_k = \begin{cases} 1 & \text{if } S_k > k\alpha + c, \qquad \text{and } S_{k+1} < (k+1)\alpha + c. \\ 0 & \text{otherwise.} \end{cases}$$

Using the same notation T_n as before, we see that T_n is a random variable which is the number of crossings of the sequence of reduced sums $S_k - k\alpha$, $k = 1, 2,$

$\cdots, n$ through the value c from above, up to possible ambiguities at $k = k_0 - 1$ or k_0 and $k = n$. Since these will not affect the value of T_n up to two units we shall ignore them in the following. We have then

$$E(Y_k) = P(S_k > k\alpha + c, S_{k+1} < (k+1)\alpha + c)$$

$$= \sum_{i=1}^{\infty} \sum_{j=-\infty}^{0} P(S_k = [k\alpha + c] + i) \tag{11}$$

$$P(X_{k+1} = [(k+1)\alpha + c] - [k\alpha + c] - i + j).$$

Now $\theta_k = [(k+1)\alpha + c] - [k\alpha + c]$ is either 0 or 1. It is easily seen that

$$k = \begin{cases} 0 & \text{if } \quad 0 < k\alpha + c - [k\alpha + c] < 1 - \alpha \\ 1 & \qquad 1 - \alpha \leqq k\alpha + c - [k\alpha + c] < 1. \end{cases}$$

Let the k's for which $\theta_k = 0$ be denoted by k_ν and these for which $\theta_k = 1$ by h_ν where both sequences $\{k_\nu\}$ and $\{h_\nu\}$ are increasing. By a theorem on uniform distribution mod 1 (Weyl [7]), we know that the sequences $\{k_\nu\}$ and $\{h_\nu\}$ have respectively the arithmetical densities $1 - \alpha$ and α.

Let $F_k(x)$ have the same meaning as before. Denoting the jump of $F_k(x)$ at ξ by $a_k(\xi)$ we have

$$E(T_n) = \sum_{k=1}^{n} \sum_{i=1}^{\infty} a_k([k\alpha + c] + i)\, F(\theta_k - i)$$

$$= \sum_{i=1}^{\infty} F(-i) \sum_{1 \leqq k_\nu \leqq n} a_{k_\nu}([k_\nu \alpha + c] + i)$$

$$+ \sum_{i=1}^{\infty} F(1 - i) \sum_{1 \leqq h_\nu \leqq n} a_{h_\nu}([h_\nu \alpha + c] + i).$$

We now use another theorem of Esseen on lattice distributions (Esseen (5), Theorem 5, p. 63), according to which[3]

$$a_k(k\alpha + \xi) = \frac{1}{(2\pi k)^{1/2}} e^{-\xi^2/2k\sigma^2} + o(k^{-1/2})$$

where $\xi = [k\alpha + c] + i - k\alpha$. For $n^{1/2-\epsilon} \leqq k \leqq n$, and values of i for which $0 \leqq \xi < k^{1/6-\epsilon}$, we have

$$\frac{\xi^2}{k} = o(1),$$

hence

$$a_k([k\alpha + c] + i) \sim \frac{1}{(2\pi k)^{1/2}}. \tag{12}$$

[3] It is interesting to remark that for any fixed ξ such that $k\alpha + \xi$ is an integer there exists an $N(\xi) > 0$ such that for all integer $k > N(\xi)$, $k\alpha + \xi$ is an actual discontinuity of $F(x)$. This follows from Esseen's result. A simple algebraical proof can be given.

We need a simple lemma, presumably known.

LEMMA. *Let $\{k_\nu\}$ be a sequence of positive integers having the arithmetical density α, viz. if we denote by $A(n)$ the number of n_i's not exceeding n, then*

$$\lim_{n\to\infty} \frac{A(n)}{n} = \alpha.$$

Let $f(x)$ be positive, monotone decreasing, and Riemann integrable in every integral $(0, n)$, and such that

$$\lim_{n\to\infty} \int_0^n f(x)dx = \infty.$$

Then as $n \to \infty$

$$\sum_{1\leq k_\nu \leq n} f(k_\nu) \sim \alpha \int_0^n f(x)dx.$$

Using the Lemma and (12), we have

$$\sum_{1\leq k_\nu\leq n} a_{k_\nu}([k_\nu\alpha + c] + i) = \sum_{1\leq k_\nu < n^{1/2-\epsilon}} + \sum_{n^{1/2-\epsilon}\leq k \leq n}$$
$$= \frac{(1-\alpha)}{\sigma}\left(\frac{2n}{\pi}\right)^{1/2} + o(n^{1/2}).$$

Similarly,

$$\sum_{1\leq h_\nu\leq n} a_{h_\nu}([h_\nu\alpha + c] + i) = \frac{\alpha}{\sigma}\left(\frac{2n}{\pi}\right)^{1/2} + o(n^{1/2}).$$

As in the case of a non-lattice distribution, the other ranges of values of k and i give a contribution of a smaller order of magnitude. Thus we obtain, from (11)

$$E(T_n) \sim \sum_{i=1}^{\infty} F(-i)\frac{(1-\alpha)}{\sigma}\left(\frac{2n}{\pi}\right)^{1/2} + \sum_{i=1}^{\infty} F(1-i)\frac{\alpha}{\sigma}\left(\frac{2n}{\pi}\right)^{1/2}.$$

We find that

$$(1-\alpha)\sum_{i=1}^{\infty} F(-i) + \alpha\sum_{i=1}^{\infty} F(1-i)$$
$$= \int_{-\infty}^{0} |x|\, dF(x) + \alpha F(0).$$

If we write

$$\gamma = \frac{\beta_1 + 2\alpha F(0)}{2\sigma} \tag{13}$$

then

$$E(T_n) \sim \gamma\left(\frac{2n}{\pi}\right)^{1/2}$$

which is the analogue of (2).

The next step is to evaluate the sum

$$\sum \prod_{\nu=1}^{l} P(S_{n_\nu - n_{\nu-1}-1} = [n_\nu \alpha] - j_{\nu-1} + i_\nu) P(S = [(n_{\nu+1})\alpha] - [n_\nu \alpha] - i_\nu + j_\nu)$$

for two fixed integers $n_0 > 0$ and j_0, when the sum runs over all possible combinations $n_1, \cdots, n_l$ between n_0 and n such that $n_i \geqq n_{i-1} + 2$ for all i. This is done by induction on l. From this point on the proof proceeds along the same lines as before.

Not let α be rational: $\alpha = r/q$, where $0 \leqq r < q$, and $(r, q) = 1$ (if $\alpha = 0$, we take $r = 0, q = 1$). Let $c = s/q + c'$ where s is an integer and $0 \leqq c' < 1/q$ Then

$$k\alpha + c = \frac{kr + s}{q} + c'.$$

Since $(r, q) = 1$, as k runs through a complete system of residues modolues q, $kr + s$ does the same. Let the k's for which

$$kr + s = 0, 1, \cdots, q - r - 1 \pmod q$$

be denoted by k_ν; and those for which

$$kr + s = q - r, \cdots, q - 1 \pmod q$$

be denoted by h_ν. The

$$0 \leqq k_\nu\alpha + c - [k_\nu\alpha + c] < 1 - \alpha$$

$$1 - \alpha \leqq h_\nu\alpha + c - [h_\nu\alpha + c] < 1,$$

and $\{k_\nu\}$ and $\{h_\nu\}$ have the arithmetical densities $(q - r)/q = 1 - \alpha$ and $r/q = \alpha$ respectively, being unions of arithmetical progressions. Define

$$Y_k = \begin{cases} 1 & \text{if } S_k > k\alpha + c, \quad S_{k+1} \leqq (k+1)\alpha + c \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$E(Y_k) = P(S_k > k\alpha + c, \quad S_{k+1} \leqq (k+1)\alpha + c)$$

$$= \sum_{i=1}^{\infty} \sum_{j=-\infty}^{0} P(S_k = [k\alpha + c] + i)$$

$$P(X_{k+1} = [(k+1)\alpha + c] - [k\alpha + c] - i + j).$$

Notice that the only difference between this and the previous case is that now it is possible to have $S_{k+1} = (k + 1)\alpha + c$ for an infinite number of values of k, if $c' = 0$ and $kr + s \equiv 0 \pmod q$. Otherwise the proof runs exactly as before.

If we define

$$Y_k = \begin{cases} 1 & \text{if } S_k \geqq k\alpha + c, \quad S_{k+1} < (k+1)\alpha + c \\ 0 & \text{otherwise,} \end{cases}$$

it can be readily verified that the same result holds. The following definition

$$Y_k = \begin{cases} 1 & \text{if } S_k > k\alpha + c, \qquad S_{k+1} < (k+1)\alpha + c \\ 0 & \text{otherwise} \end{cases}$$

however, leads to a different result. We may also include the possibility that $S_{k-1} > (k-1)\alpha + c$, $S_k = k\alpha + c$, $S_{k+1} < (k+1)\alpha + c$ and other variations. They can all be treated by the method developed here and offer no novelty.

Keeping these alternatives in mind we state our result as follows:

THEOREM 2. *Suppose that $F(x)$ is of the lattice type with span* 1 *and such that* $0 \leqq \alpha < 1$, $\sigma^2 > 0$, $\beta_3 < \infty$. *Let c be a given real number and let T_n be the number of k's not exceeding n for which*

$$S_k - k\alpha > c, \qquad S_{k+1} - (k+1)\alpha \leqq c.$$

Then

$$\lim_{n\to\infty} P(T_n \leqq \gamma n^{1/2}x) = \left(\frac{2}{\pi}\right)^{1/2} \int_0^x e^{-y^2/2}\,dy$$

where

$$\gamma = \tfrac{1}{2}(\beta_1 + 2\alpha F(0)).$$

CORNELL UNIVERSITY

REFERENCES

(1) PAUL LÉVY, *Sur certains processus stochastiques homogènes*, Compositio Math. vol. 7 (1939), pp. 283–339.
(2) W. FELLER, *The fundamental limit-theorems in probability*, Bull. Amer. Math. Soc. vol. 51 (1945), pp. 800–832.
(3) W. FELLER, *Fluctuation theory of recurrent events*, Trans. Amer. Math. Soc. vol. 67. (1949), pp. 98–119.
(4) KAI LAI CHUNG AND G. A. HUNT, *On the zeros of* $\sum \pm 1$, Ann. of Math., vol. 50, (1949), pp. 385–400.
(5) C. G. ESSEEN, *Fourier analysis of distribution functions*, Acta Math. vol. 77 (1945), pp. 5–125.
(6) G. PÓLYA, *Ueber den zentralen Grenzweitsatz der Wahrscheinlichkeitsrechung*, Math. Zeit. vol. 8 (1920), pp. 171–181.
(7) H. WEYL, *Ueber die Gleichveiteilung von Zahlen mod. Eins*, Math. Ann. vol. 77 (1916) pp. 313–352.

ANNALS OF MATHEMATICS
Vol. 57, No. 3, May, 1953
Printed in U.S.A.

CORRECTIONS TO MY PAPER "FLUCTUATIONS OF SUMS OF INDEPENDENT RANDOM VARIABLES"

K. L. CHUNG

(Received October 7, 1952)

The following misprints (omitting several obvious ones) and errors in the paper named in the title (Ann. of Math., Vol. 51, (1950) 697–706) have been found by Mr. L. J. Cote:

1. In the line above (1), insert "if $|x| k^{-\frac{1}{2}} = o(1)$" before "It follows that". In (1), the error term $o(k^{-\frac{1}{2}})$ should be $o((|x| + 1)k^{-\frac{1}{2}})$.

2. The right side of (1 bis) should be

$$\frac{|x|}{(2\pi k)^{\frac{1}{2}}} + o\left(\frac{|x| + 1}{k^{\frac{1}{2}}}\right).$$

The o-term in the last line on p. 698 should be

$$o\left(\sum_{k=1}^{n} \frac{|x| + 1}{k^{\frac{1}{2}}}\right).$$

3. In the following places the exponent $\frac{1}{2}$ should be $l/2$: (6), line 3 and 4 from bottom on p. 700, line 3 from bottom on p. 701.

4. In the last display on p. 701, the second half of the formula should precede the first half.

5. On p. 700, line 11–12, $W(n_1, n, y_0, l)$ should be $W(n_1, n, y_1, l)$. This necessitates a modification of the arguments leading from (6) to (7) on pp. 700–701. The correct argument is sketched as follows.

We write

$$\text{(I)}\qquad \begin{aligned} W(n_0, n, y_0, l + 1) &= \sum_{n_0 < n_1 < n} \int_{x_1=0+}^{\infty} dF_{n_1-n_0-1}(-y_0 + x_1) \\ &\quad \cdot \int_{y_1=-\infty}^{0-} dF(-x_1 + y_1) W(n_1, n, y_1, l) \\ &= \sum_{n_0 < n_1 < n} \iint_{\Delta} W(n_1, n, y_1, l)\, dH_{n_1-n_0}(x_1, y_1). \end{aligned}$$

The induction hypothesis is two-fold:

$$\text{(IIa)}\qquad W(n_0, n, y_0, l) \leq A_l(n - n_0)^{l/2}$$

where A_l is a constant which depends only on $F(x)$ and l; and, for a fixed y_0, as $n - n_0 \to \infty$

$$\text{(IIb)}\qquad W(n_0, n, y_0, l) \sim c_l(n - n_0)^{l/2}$$

uniformly in n_0. In fact, W as a function of n_0 and n depends only on $n - n_0$.

That (IIa) and (IIb) are true for $l = 1$ is proved by the arguments on pp. 698–699, noticing that

$$|F_k(-y_0 - x) - F_k(-y_0)| \leqq A(|x| + 1)k^{-\frac{1}{2}}$$

where A depends only on $F(x)$.

From (1 bis) corrected it follows that as $n_1 - n_0 \to \infty$

$$\text{(III)}\quad \iint_\Delta dH_{n_1-n_0}(x_1, y_1) = \int_{-\infty}^{0-} \{F_{n_1-n_0-1}(-y_0 - x) - F_{n_1-n_0-1}(-y_0)\}\, dF(x)$$

$$\sim \beta_1\, 2^{-\frac{1}{2}}\pi^{-\frac{1}{2}}(n_1 - n_0)^{-\frac{1}{2}}.$$

By virtue of (IIa) we can apply Lebesgue's convergence theorem and obtain, as $n - n_1 \to \infty$, uniformly with respect to n_1,

$$\text{(IV)}\quad \iint_\Delta W(n_1, n, y_1, l)\, dH_{n_1-n_0}(x_1, y_1) \sim c_l\, (n - n_1)^{l/2} \iint_\Delta dH_{n_1-n_0}(x_1, y_1).$$

From (I), (III), (IV) and (IIa), it follows by elementary limiting processes that for a fixed y_0 as $n - n_0 \to \infty$

$$W(n_0, n, y_0, l + 1) \sim \sum_{n_1=n_0+1}^{n-1} \frac{\beta_1 c}{2(2\pi)^{\frac{1}{2}}} (n_1 - n_0)^{-\frac{1}{2}}(n - n_1)^{l/2}.$$

Thus (IIb) is true with $l + 1$ replacing l, where c_{l+1} is given by (7).

From (I), (IIa) and (III) it follows that

$$W(n_0, n, y_0, l + 1) \leqq A_l A' \textstyle\sum_{n_0<n<n}(n - n_1)^{l/2}(n_1 - n_0)^{-\frac{1}{2}}$$

where A' depends only on $F(x)$. Hence (IIa) is true with $l + 1$ replacing l. The induction is complete.

CORNELL UNIVERSITY

Reprinted with kind permission from *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950*, University of California Press, Berkeley and Los Angeles (1951), 341–352.

THE STRONG LAW OF LARGE NUMBERS

KAI LAI CHUNG
CORNELL UNIVERSITY

1. Introduction

A well known unsolved problem in the theory of probability is to find a set of necessary and sufficient conditions (nasc's) for the validity of the strong law of large numbers (SLLN) for a sequence of independent random variables. This problem will not be solved in the present paper. To avoid a possible misunderstanding it must be stated at once that nasc's have been found, and several sets of them will be given in section 3, but they are all unsatisfactory. Presumably all (or shall we say most) mathematicians will agree on a satisfactory set of such conditions if and when they are exhibited, but before they are it does not seem easy to lay down criteria of satisfactoriness. On the other hand it is safe to rule out certain conditions as unsatisfactory, for example those in which sums of random variables enter; the conditions to be given in section 3 all have this undesirable property.

The purpose of this paper is to give an account of the latest information on this problem, at least in some directions. While undoubtedly much that follows is known to experts in the field or, so to speak, lurks in the corners of their minds, it is hoped that some of the results below are printed here for the first time and not sufficiently known to a wider circle of probabilists. It is to acquaint this latter group with the present status of knowledge of the problem that this paper is written.

The paper is divided into three sections. Section 2 is quite independent of the others and deals with the case of identically distributed, independent random variables (r.v.'s). In this case it is known, after Kolmogorov,[1] that a nasc for the validity of the SLLN is the finiteness of the first absolute moment of the common distribution function (d.f.). For use in certain statistical applications Professor Wald raised the question of the uniformity of the strong convergence with respect to a family of d.f.'s (see section 2). A nasc for this is given in section 2, which includes Kolmogorov's theorem as a special case. The method of proof is classical.

In section 3 several sets of nas, but unsatisfactory, conditions for the validity of the SLLN are given and their interrelations, mostly trivial, are explored. The results of this section includes Kawata's partial result[2] in this direction, and Pro-

This research was prepared at the University of California under the partial support of the Office of Naval Research.

[1] See Kolmogorov [1, p. 67]. As far as the author is aware the proof was never published by him. The proof of the sufficiency part is given in Fréchet [2]. The necessity part has been given without centering at the medians; see Feller [3], for more general results.

[2] Kawata [4]. He stated the theorem with zero expectations, an assumption which he never used.

horov's result recently announced [5].* The proof given here of Prohorov's result is different from and somewhat longer than his,[3] but it is hoped that it brings out the connections more clearly. As an application a simple proof of a sufficient condition which includes Kolmogorov's [7] and Brunk's [8] is given, as also announced by Prohorov.

In section 4 satisfactory nasc's for the SLLN are found for r.v.'s which are individually bounded and whose bounds satisfy certain restrictive order conditions. Such a result was also announced by Prohorov. By using a deep estimate due to Cramér and Feller [9], Prohorov's result is extended to slightly more general cases

In the following $\{X_n\}$, $n = 1, 2, \ldots$ will always denote a sequence of independent real valued r.v.'s, and $S_n = \sum_{k=1}^{n} X_k$. If X is a r.v., $m(X)$ denotes a median[4] of X; X^0 denotes the centered r.v. $X - m(X)$; $E(X)$ the expectation of X. If A is an event, $P(A)$ denotes its probability. The letters "i.o." are an abbreviation of the phrase "infinitely often," namely, "for an infinite number of values of whatever subscript is in question." The symbol ϵ denotes an arbitrarily small positive number, thus a proposition involving ϵ should read: "For every $\epsilon > 0$, etc."

If there exists a sequence of real numbers $\{c_n\}$ such that

$$P\left(\lim_{n\to\infty} \frac{S_n - c_n}{n} = 0\right) = 1 \tag{1.1}$$

we say that the sequence $\{X_n\}$ obeys the SLLN. In this case it is trivial that we can replace c_n by $m(S_n)$. Thus (1.1) is equivalent to

$$P\left(\lim_{n\to\infty} \frac{S_n^0}{n} = 0\right) = 1 \tag{1.1 bis}$$

or to the following:

$$P(|S_n^0| > n\epsilon \text{ i.o.}) = 0 \tag{1.2}$$

or to

$$P(|S_n^0| \leqq n\epsilon \quad \text{for all} \quad n \geqq N) = 1\,. \tag{1.3}$$

Note that in (1.3) the N is allowed to depend not only on ϵ, but also on the sample sequence $\{X_n\}$. Thus (1.3) is equivalent to the following: given any $\epsilon > 0$, there exists a fixed N_0 depending on ϵ but no longer on the sample sequence such that

$$P(|S_n^0| \leqq n\epsilon \quad \text{for all} \quad n \geqq N_0) \geqq 1 - \epsilon\,. \tag{1.4}$$

2. The identically distributed case

Let all X_n have the same d.f. $F(x)$. Kolmogorov (see footnote 2) proved that

* *Added in proof:* Prohorov's complete account has in the meanwhile appeared in *Izvestia Akad. Nauk. USSR*, Vol. 14 (1950), pp. 523–536.

[3] His proof depends on a new inequality of Kolmogorov, the idea of which is very close to one of P. Lévy [6, p. 138].

[4] Throughout this paper we could use instead of the median, any number $\mu(X)$ such that $P[X \geqq \mu(X)] \geqq \lambda$, $P[X \leqq \mu(X)] \geqq \lambda$ for some fixed λ: $0 < \lambda < 1$.

a nasc that $\{X_n\}$ obeys the SLLN is that

$$\int_{-\infty}^{\infty} |x|\, dF(x) < \infty .$$

The proof of the necessity part is trivial; as to the proof of the sufficiency part there are three essentially different methods:

(i) Khintchine-Kolmogorov's method which depends on truncation and Kolmogorov's famous inequality with or without the intervention of infinite series;

(ii) A special case of G. D. Birkhoff's individual ergodic theorem, ever so many proofs of which have been given;[5]

(iii) Doob's [11] very elegant proof using the theory of martingales.

The proof of the following more general theorem uses method (i) and is in essence nothing but a precision of that method. It is not clear whether the other methods will be applicable.

Let a family of d.f.'s $F(x, \theta)$ be given where θ is the parameter of the family. All the r.v.'s X_n have one and the same d.f. $F(x, \theta)$ from the family where θ may be any value of the parameter. The sequence $\{X_n\}$ is said to obey the SLLN uniformly with respect to θ if: given any $\epsilon > 0$, there exists a fixed $N_0 = N_0(\epsilon)$ not depending on θ such that (1.4) holds no matter what θ is.

THEOREM. *A sufficient condition for the sequence $\{X_n\}$ to obey the SLLN uniformly with respect to θ is the following: given any δ there exists a number $A(\delta)$ not depending on θ such that*

$$\int_{|x|>A(\delta)} |x|\, dF(x, \theta) < \delta . \tag{2.1}$$

If so we can replace S_n^0 in (1.4) by $S_n - E(S_n)$. This condition is also necessary if the median $m(\theta)$ of $F(x, \theta)$ is a bounded function of θ.

PROOF. *Sufficiency*. From (2.1) it follows that

$$\int_{-\infty}^{\infty} |x|\, dF(x, \theta) \leqq A(1) + 1 = M .$$

Now choose $N \geqq 2$ and such that

$$\int_{|x|\geqq N} |x|\, dF(x, \theta) + \frac{16}{\epsilon^2}\left(\frac{2}{N^{1/2}} + 6 \int_{|x|\geqq N^{1/4}} |x|\, dF(x, \theta)\right) < \frac{\epsilon}{4}. \tag{2.2}$$

Having chosen N, choose $N_0 > N$ such that

$$\frac{16N^2}{N_0\epsilon} \int_{|x|>N_0\epsilon/8N} |x|\, dF(x, \theta) + \frac{NM}{N_0} < \frac{\epsilon}{4}. \tag{2.3}$$

We have

$$\sum_{k=N}^{\infty} P(|X_k| \geqq k) \leqq \int_{|x|\geqq N} |x|\, dF(x, \theta). \tag{2.4}$$

[5] Extensions of methods (i) and (ii) to the case of dependent r.v.'s which includes the case of independence have been announced by M. Loève [10].

Next,

$$(2.5)\quad \sum_{k=N}^{\infty}\frac{1}{k^2}\int_{|x|<k}|x|^2 dF(x,\theta)=\sum_{k=0}^{\infty}\int_{k\leq|x|<k+1}|x|^2 dF(x,\theta)\sum_{j=\max(N,k)}^{\infty}\frac{1}{j^2}$$

$$\leq\frac{1}{N-1}\int_{0\leq|x|<N^{1/4}}|x|^2 dF(x,\theta)+\frac{N+1}{N-1}\int_{|x|\geq N^{1/4}}|x|\,dF(x,\theta)$$

$$+\frac{N+2}{N}\int_{|x|\geq N+1}|x|\,dF(x,\theta)\leq\frac{2}{N^{1/2}}+6\int_{|x|\geq N^{1/4}}|x|\,dF(x,\theta).$$

Define

$$X_k'=\begin{cases}X_k & \text{if}\quad |X_k|<k,\\ 0 & \text{if}\quad |X_k|\geq k\end{cases}$$

$$X_k''=X_k'-E(X_k').$$

Then

$$E(X_k''^2)\leq\int_{|x|<k}|x|^2 dF(x,\theta).$$

By Kolmogorov's inequality, and (2.3)-(2.5),

$$P\left(\left|\sum_{k=N}^{n}\frac{X_k-E(X_k')}{k}\right|>\frac{\epsilon}{4}\text{ for at least one }n\geq N\right)<\frac{\epsilon}{4}.$$

It follows from Kronecker's lemma that

$$P\left(\left|\frac{1}{n}\sum_{k=N}^{n}[X_k-E(X_k')]\right|>\frac{\epsilon}{4}\text{ for at least one }n\geq N\right)<\frac{\epsilon}{4}.$$

Moreover, we have

$$P\left(\left|\frac{X_1+\ldots+X_N}{n}\right|>\frac{\epsilon}{4}\right)\leq N\int_{|x|>n\epsilon/4N}dF(x,\theta)\leq\frac{16N^2}{n\epsilon}$$

$$\times\int_{|x|>n\epsilon/8N}|x|\,dF(x,\theta)<\frac{\epsilon}{4},$$

and if $n\geq N_0$, by (2.2) and (2.3)

$$\left|\frac{1}{n}\sum_{k=N}^{n}E(X_k')\right|\leq\frac{NM}{n}+\int_{|x|\geq N}|x|\,dF(x,\theta)<\frac{\epsilon}{2}.$$

Altogether we conclude that

$$P\left(\left|\frac{S_n-E(S_n)}{n}\right|>\epsilon\text{ for at least one }n\geq N_0\right)<\frac{\epsilon}{2},$$

or

$$P[|S_n-E(S_n)|\leq n\epsilon\text{ for all }n\geq N_0]\geq 1-\epsilon.$$

If $\epsilon<\frac{1}{2}$ it follows by the definition of a median that for all $n\geq N_0$,

$$(2.6)\qquad |m(S_n)-E(S_n)|\leq n\epsilon.$$

Thus

$$P\,[|S_n - m(S_n)| \leqq 2n\epsilon \quad \text{for all} \quad n \geqq N_0] \geqq 1 - \epsilon\,.$$

This implies (1.4), whatever θ is.

Necessity. Suppose that (1.4) holds where N_0 does not depend on θ. Then if $n \geqq N_0$,

$$P\,[|X_n - m(S_n) + m(S_{n-1})| \leqq 2n\epsilon] \geqq 1 - \epsilon\,.$$

If $\epsilon < \frac{1}{2}$ this entails

$$|m(X_n) + m(S_{n-1}) - m(S_n)| \leqq 2n\epsilon\,.$$

Since by hypothesis $|m(X_n)| = |m(\theta)| \leqq m$ where m does not depend on θ, there exists a number N_1, not depending on θ, such that if $n \geqq N_1$

$$|m(S_{n-1}) - m(S_n)| < 3n\epsilon\,. \tag{2.7}$$

Now suppose that (2.1) did not hold and we wish to reach a contradiction. If (2.1) did not hold there exists a $\delta > 0$ such that for any N there is a θ_N for which

$$\int_{|x|>N} |x|\,dF(x, \theta_N) \geqq \delta\,.$$

Hence

$$\sum_{k=N+1}^{\infty} \int_{|x|>k} dF(x, \theta_N) + (N+1) \int_{|x|>N} dF(x, \theta) \geqq \delta\,.$$

It follows that one of the following two cases would occur:

Case (i). For a sequence $N_i \uparrow \infty$, there corresponds a sequence θ_i such that if all X_n have the d.f. $F(x, \theta_i)$, then

$$\sum_{n=N_i+1}^{\infty} P(|X_n| \geqq n) \geqq \frac{\delta}{2}.$$

Hence for this sequence $\{X_n\}$ we have

$$\sum_{n=N_i+1}^{\infty} P(|X_n| < n) \leqq e^{-\delta/2}.$$

$$P(|X_n| \geqq n \quad \text{for at least one} \quad n \geqq N_i + 1) \geqq 1 - e^{-\delta/2}\,. \tag{2.8}$$

We have $X_n = S_n - m(S_n) - [S_{n-1} - m(S_{n-1})] + m(S_n) - m(S_{n-1})$; by (2.7), since $N_i \geqq N_1$, if $\epsilon < 1/6$, (2.8) entails

$$P\,[|S_n - m(S_n)| \geqq \frac{n}{4} \quad \text{for at least one} \quad n \geqq N_i] \geqq 1 - e^{-\delta/2}\,.$$

Since $N_i \uparrow \infty$, (1.4) becomes false for $\epsilon < \min(1/4, 1 - e^{-\delta/2})$.

Case (ii). For a sequence $N_i' \uparrow \infty$, there corresponds a sequence θ_i' such that if all X_n have the d.f. $F(x, \theta_i')$, then

$$\sum_{n=N_i'}^{2N_i'} P(|X_n| > N_i') \geqq \frac{\delta}{2}$$

whence

$$P(|X_n| > N_i' \quad \text{for at least one } n: N_i' \leq n \leq 2N_i') \geqq 1 - e^{-\delta/2}.$$

The same argument as in case (i) finishes the proof.

Remark 1. If the family of d.f.'s consists of a single d.f., then the theorem reduces to Kolmogorov's.

Remark 2. Without the assumption of the boundedness of the medians, the condition stated in the theorem is not necessary.

Example. Let θ run over the positive integers and define $F(x, n)$ to be the d.f. which has a single jump at the point $x = n$.

Remark 3. The following simpler version may be more useful for applications; its proof is similar but simpler. Suppose that for every θ,

$$\int_{-\infty}^{\infty} x\,dF(x, \theta) = 0, \qquad \int_{-\infty}^{\infty} |x|\,dF(x, \theta) < \infty.$$

Then the condition stated in the theorem is a nasc that: given any $\epsilon > 0$, there exists a N_0 depending on ϵ but not on θ nor on the sample sequence, such that

$$P(|S_n| \leq n\epsilon \quad \text{for all} \quad n \geq N_0) \geqq 1 - \epsilon.$$

3. Necessary and sufficient conditions and a sufficient condition

We return now to the general case. We shall consider, besides the SLLN embodied in formula (1.1), also a modified form, namely,

$$P\left(\lim_{n\to\infty} \frac{S_n}{n} = 0\right) = 1. \tag{3.1}$$

For any given sequence c_n, (1.1) can be reduced to (3.1) by an obvious change of variables: $X_n^* = X_n - c_n + c_{n-1}$. Thus while (1.1) answers the question: does there exist some sequence $\{c_n\}$ such that (1.1) holds; (3.1) answers the question: does (1.1) hold with a given sequence c_n. The second question will of course be answered via the first if we can decide whether $\lim_{n\to\infty} \frac{m(S_n) - c_n}{n} = 0$ or not, but there seems in general no control over $m(S_n)$.

To simplify writing, "convergence in probability" will be denoted by an arrow $\to$; "convergence with probability one" or "almost sure convergence" by a double arrow $\rightrightarrows$. If A and B are two propositions, $A \supset B$ means "A implies B"; $A \equiv B$ means "A and B are equivalent."

Consider the following:

$$\frac{S_n}{n} \rightrightarrows 0; \tag{1}$$

$$\frac{S_n}{n} \to 0; \tag{2}$$

$$\frac{S_{2^n}}{2^n} \rightrightarrows 0;[6] \tag{3}$$

[6] Instead of 2^n we can take any sequence of positive integers such that $0 < A_1 < q_{n+1}/q_n < A_2 < \infty$.

$$\frac{S_{2^{n+1}} - S_{2^n}}{2^n} \rightrightarrows 0\,; \tag{4}$$

$$\sum_n P\left(|S_{2^{n+1}} - S_{2^n}| > 2^n\epsilon\right) < \infty\,; \tag{5}$$

$$\sum_n P\left(|S_{2^n}| > 2^n\epsilon\right) < \infty\,. \tag{6}$$

The following relations obtain:

I. (1) $\supset$ (2): Well known.

II. (1) $\supset$ (3) $\equiv$ (4) $\equiv$ (5): The equivalence of (3) and (4) is a simple analytical fact; that of (4) and (5) is a consequence of the Borel-Cantelli lemma.

III. (5) $\subset$ (6) $\supset$ (3): That (6) implies (3) is a consequence of one half of the Borel-Cantelli lemma; that (6) implies (5) follows from Boole's inequality.

IV. (2) and (3) $\supset$ (1).

PROOF. (3) is equivalent to

$$P\left(|S_{2^n}| > 2^n\epsilon \text{ i.o.}\right) = 0\,.$$

For every positive integer k define $n(k)$ by $2^{n(k)-1} \leqq k < 2^{n(k)}$. Since

$$P\left(|S_{2^{n(k)}} - S_k| \leqq 2^{n(k)}\epsilon\right) \geqq 1 - P\left(|S_{2^{n(k)}}| > 2^{n(k)-1}\epsilon\right) - P\left(|S_k| > 2^{n(k)-1}\epsilon\right)$$

by (2) we have if $k > k_0 = 2^{n_0}$

$$P\left(|S_{2^{n(k)}} - S_k| \leqq 2^{n(k)}\epsilon\right) > \tfrac{1}{2}\,.$$

Now $|S_k| > 2^{n(k)+1}\epsilon$ and $|S_{2^{n(k)}} - S_k| \leqq 2^{n(k)}\epsilon$ together imply $|S_{2^{n(k)}}| > 2^{n(k)}\epsilon$ and the first two events are independent, hence by a simple argument,

$$P\left(|S_k| > 2^{n(k)+1}\epsilon \text{ for some } k > k_0\right) \leqq 2P\left(|S_{2^{n(k)}}| > 2^{n(k)}\epsilon \text{ for some } k > k_0\right).$$

Letting $k_0 \to \infty$ we obtain

$$P\left(|S_k| > 4k\epsilon \text{ i.o.}\right) \leqq 2P\left(|S_{2^n}| > 2^n\epsilon \text{ i.o.}\right) = 0\,.$$

V. (1) $\equiv$ (2) and (3) $\equiv$ (2) and (4) $\equiv$ (2) and (5) $\subset$ (2) and (6): from I–IV.

The implication (2) and (6) $\supset$ (1) was proved by Kawata [4]. Proposition (2) is one form of the weak law of large numbers (WLLN). Thus the relations in V show that the SLLN, in the form (2.1), is equivalent to the corresponding WLLN plus the SLLN for the subsequence S_{2^n}. Now satisfactory nasc for the WLLN have been given by Kolmogorov [12] and Feller [13]. Hence we can, if we prefer, replace (2) everywhere in V by these conditions. The significance of Prohorov's result below lies in the elimination of the WLLN as part of the sufficient conditions for the SLLN, and this is done by centering (at the medians).

Now we consider the relations (1)–(6) with S_n, S_{2^n}, $S_{2^{n+1}} - S_{2^n}$ replaced by S_n^0, $S_{2^n}^0$, $(S_{2^{n+1}} - S_{2^n})^0$ respectively. (Note that $S_{2^{n+1}} - S_{2^n}$ is *not* replaced by $S_{2^{n+1}}^0 - S_{2^n}^0$!). The resulting propositions we call (1^0)–(6^0). We add a new proposition

$$\frac{S_{2^n}^0}{2^n} \to 0\,. \tag{7^0}$$

We notice first that (7^0) entails

$$\frac{S_{2^{n+1}} - S_{2^n} - m(S_{2^{n+1}}) - m(S_{2^n})}{2^n} \to 0;$$

hence it follows that

$$(8^0) \qquad \frac{m(S_{2^{n+1}}) - m(S_{2^n}) - m(S_{2^{n+1}} - S_{2^n})}{2^n} \to 0.$$

All the relations I–III above carry over for the circled propositions and the corresponding relations will be referred to as I^0–III^0. We have only to consider $X_n - m(S_n) + m(S_{n-1})$ as new r.v.'s and apply I–III; the fact that (7^0) $\supset$ (8^0) has to be used in several places.

IV^0. (3^0) $\supset$ (7^0) $\supset$ (2^0).

Proof. Only the second implication needs a proof and this is easiest done by resorting to characteristic functions (c.f.'s). Let X_n have the c.f. $f_n(t)$. (7^0) implies

$$\lim_{n\to\infty} \prod_{j=1}^{2^n} \left| f_j\left(\frac{t}{2^n}\right)\right| = 1$$

uniformly in any finite interval $|t| \leq T$. This immediately implies

$$\lim_{k\to\infty} \prod_{j=1}^{k} \left| f_j\left(\frac{2^{n(k)}}{k}\cdot\frac{t}{2^{n(k)}}\right)\right| = 1$$

uniformly in $|t| \leq T/2$, where $n(k)$ is defined in IV, whence (2^0).

V^0. (5^0) $\supset$ (1^0).

Proof. (5^0) $\equiv$ (3^0) $\supset$ (7^0) $\supset$ (8^0). Since (8^0) holds (5^0) is equivalent to

$$(9^0) \qquad \sum P(|S^0_{2^{n+1}} - S^0_{2^n}| > 2^n\epsilon) < \infty.$$

Moreover by II^0 and IV^0, (5^0) $\supset$ (2^0). But (2^0) and (9^0) imply (1^0), by the third proposition in V.

VI^0. (1^0) $\equiv$ (3^0) $\equiv$ (4^0) $\equiv$ (5^0) $\subset$ (6^0): from II^0, III^0, IV^0, V^0.

The equivalence (1^0) $\equiv$ (5^0) is Prohorov's theorem [5].

We shall now prove the following theorem which gives a sufficient condition for the SLLN and includes Kolmogorov's ($r = 1$) and Brunk's (r integer $\geqq 1$).

Theorem. *Let $E(X_n) = 0$ for every n, and $E(|X_n|^{2r}) < \infty$ for some real number $r \geqq 1$. If*

$$(3.2) \qquad \sum_n \frac{E(|X_n|^{2r})}{n^{r+1}} < \infty.$$

Then (2.1) *holds.*

Proof. We need the following inequality

$$(3.3) \qquad E\left(\left|\sum_{k=1}^n X_k\right|^{2r}\right) \leqq A n^{r-1} \sum_{k=1}^n E(|X_k|^{2r})$$

where A depends only on r. This is easily proved if we use an inequality due to Marcinkiewicz and Zygmund [14] (trivial if r is an integer) according to which

$$E\left(\left|\sum_{k=1}^n X_k\right|^{2r}\right) \leqq A E\left[\left(\sum_{k=1}^n X_k^2\right)^r\right].$$

Now by Hölder's inequality

$$\left(\sum_{k=1}^{n} X_k^2\right)^r \leq n^{r-1} \sum_{k=1}^{n} |X_k|^{2r};$$

hence (3.3) follows.

From (3.2) it follows firstly by Kronecker's lemma

$$\overline{\lim_{n\to\infty}} E\left(\left|\frac{S_n}{n}\right|^{2r}\right) \leq \lim_{n\to\infty} \frac{A}{n^{r+1}} \sum_{k=1}^{n} E(|X_k|^{2r}) = 0.$$

Hence by Tschebicheff's inequality, we have proposition (2) above. Next, again by Tschebicheff's inequality, and (3.3)

$$P(|S_{2^{n+1}} - S_{2^n}| > 2^n\epsilon) \leq \frac{A\, 2^{n(r-1)}}{(2^n\epsilon)^{2r}} \sum_{k=2^n+1}^{2^{n+1}} E(|X_k|^{2r}) \leq \frac{A\, 2^{r+1}}{\epsilon^{2r}} \sum_{k=2^n+1}^{2^{n+1}} \frac{E(|X_k|^{2r})}{k^{r+1}}.$$

Hence from (3.2) follows proposition (5). Since (2) and (5) $\equiv$ (1) by VI we have (1).

Remark. Using truncated variables the theorem can be stated without assuming any moments. We shall not insist on this, and also other more or less trivial extensions of the theorem [15].

4. Necessary and sufficient conditions for some special cases

For easier reference we shall rewrite some of the previous formulas:

$$\frac{S_n^0}{n} \rightrightarrows 0, \tag{4.1}$$

$$\frac{S_n}{n} \rightrightarrows 0. \tag{4.2}$$

If (4.1) holds, it is easy to see that we have

$$\sum P(|X_n^0| \geq n\epsilon) < \infty. \tag{4.3}$$

Define $X_n' = X_n^0$ if $|X_n^0| \leq n\epsilon$, and $X_n' = 0$ if $|X_n^0| > n\epsilon$; then under (4.3) the sequences $\{X_n'\}$ and $\{X_n^0\}$ are equivalent in the sense of Khintchine. If the SLLN is valid for $\{X_n\}$ it is by definition also valid for $\{X_n^0\}$ and so for $\{X_n'\}$; conversely if the SLLN is valid for $\{X_n'\}$, and (4.3) is assumed, then it is also valid for $\{X_n\}$. Hence we may confine ourselves to r.v.'s $\{X_n\}$ satisfying the following condition:

$$\sup |X_n| = o(n). \tag{4.4}$$

Under (4.4) we may, without loss of generality, assume that

$$E(X_n) = 0, E(X_n^2) = \sigma_n^2, \sum_{k=1}^{n} \sigma_k^2 = s_n^2 [= s^2(n)], s^2(2^{n+1}) - s^2(2^n) = d_n^2. \tag{4.5}$$

LEMMA. *Under* (4.4) *and* (4.5), (4.1) *and* (4.2) *are equivalent.*

PROOF. We need only prove that (4.1) implies (4.2). We shall first prove that (4.1) implies $s_n = o(n)$. Otherwise let $n_k \uparrow \infty$ and $s_{n_k} \geq \delta n_k$ for all k with $\delta > 0$.

Then by (4.4) we have $\max_{1 \leq j \leq n_k} \sup |X_j| = o(s_{n_k})$. By a classical theorem of P. Lévy ([6], p. 102) the central limit theorem holds for the sequence S_{n_k}, in particular for every $\eta > 0$

$$\lim_{k \to \infty} P(S_{n_k} \geqq \eta s_{n_k}) = \lim_{k \to \infty} P(S_{n_k} \leqq -\eta s_{n_k}) = \frac{1}{\sqrt{2\pi}} \int_\eta^\infty e^{-y^2/2} dy < \tfrac{1}{2}.$$

It follows that $m(S_{n_k}) = o(s_{n_k})$. Consequently the d.f. of $s_{n_k}^{-1}(S_{n_k}^0)$ tends to that of the normal and (4.1) cannot be true.

Therefore $s_n = o(n)$ and it follows that $n^{-1}S_n \to 0$ (in probability !). This and (4.1) imply that $m(S_m) = o(n)$ and hence (4.2). q.e.d.

Of the results in section 3 we shall use the following which, combined with the lemma above, will be referred to as (P).

(P) Under (4.4) and (4.5): if $s_n = o(n)$ and also

$$(4.6) \qquad \sum P(|S_{2^{n+1}} - S_{2^n}| > 2^n \epsilon) < \infty$$

then both (4.1) and (4.2) hold; if (4.1) or (4.2) holds, then (4.6) holds.

In the following (4.4) and (4.5) will be assumed.

THEOREM 1. *If the further condition is satisfied:*

$$(4.7) \qquad M_n = \max_{2^n < k \leqq 2^{n+1}} \sup |X_k| = o\left(\frac{d_n^2}{2^n}\right),$$

then a nasc for (4.1) *or* (4.2) *is*

$$(4.8) \qquad \sum \exp(-\epsilon 2^{2n} d_n^{-2}) < \infty .$$

PROOF. *Sufficiency.* If (4.8) holds, then $s_n = o(n)$ because s_n is nondecreasing. Furthermore, by a theorem of Feller and using his notation, see [9],

$$\max_{2^n < k \leqq 2^{n+1}} \sup |X_k| \leqq \lambda_n d_n , \quad \lambda_n = 2^{-n} d_n = o(1), \quad x = \epsilon 2^n d_n^{-1}, \quad 0 < \lambda_n x = \epsilon$$

we have

$$(4.9) \qquad P(|S_{2^{n+1}} - S_{2^n}| > 2^n \epsilon) \sim \frac{C d_n}{2^n \epsilon} \exp\{(-\epsilon^2 2^{2n} d_n^{-2})(1 + \delta)\}$$

where $\delta \to 0$ as $\epsilon \to 0$, and C is an absolute constant. Thus (4.8) implies (4.6). (4.2) and (4.1) follow by (P).

Necessity. If (4.1) or (4.2) holds, (4.6) holds by (P). Also from the proof of the lemma, we have $s_n = o(n)$. Hence the estimate (4.9) is applicable and (4.6) reduces to (4.8).

Remark. Obviously (4.8) is a nasc for (4.1) or (4.2) if all the X_n have a normal distribution with zero mean.

THEOREM 2. *If instead of* (4.7) *we assume*

$$(4.10) \qquad \sup |X_n| = o\left(\frac{n}{\lg \lg n}\right)$$

then (4.8) *is again a nasc for* (4.1) *or* (4.2).

PROOF. It follows from (4.10) that $M_n = o(2^n \lg^{-1} n)$. Let the values of n for

which $M_n > 2^{-n}d_n^2$ be n_k, $k = 1, 2, \ldots$. Let $M_n^* = \epsilon^{-1}M_n(> M_n$ if $\epsilon < 1)$. Then

$$\frac{d_{n_k}^2}{\epsilon 2^{n_k}} < M_{n_k}^* = o\left(\frac{2^{n_k}}{\lg n_k}\right).$$

By an inequality of Kolmogorov [16], we have if $n = n_k$

$$P(|S_{2^{n+1}} - S_{2^n}| > 2^n\epsilon) \leqq \max\left[\exp\left(\frac{-2^n\epsilon}{4M_n^*}\right), \exp\left(-\frac{2^{2n}\epsilon^2}{4d_n^2}\right)\right] = \exp\left(-\frac{2^n\epsilon}{4M_n^*}\right). \tag{4.11}$$

It follows from (4.11) that

$$\sum_k P(|S_{2^{n_k+1}} - S_{2^{n_k}}| > 2^{n_k}\epsilon) < \infty . \tag{4.12}$$

On the other hand it is trivial that

$$\sum_k \exp[-\epsilon 2^{2n_k} d^{-2}(n_k)] < \infty . \tag{4.13}$$

Now, if (4.1) or (4.2) holds, then (4.6) holds by (P). If $n \neq n_k$, (4.9) is applicable, hence

$$\sum_{n \neq n_k} \exp[-\epsilon 2^{2n} d^{-2}(n)] < \infty .$$

This and (4.13) give (4.8).

Conversely, if (4.8) holds, then by (4.9)

$$\sum_{n \neq n_k} P(|S_{2^{n+1}} - S_{2^n}| > 2^n\epsilon) < \infty .$$

This and (4.12) give (4.6). (4.1) and (4.2) follow by (P).

Theorem 2 was announced by Prohorov. If s_n is of a greater order of magnitude than $n(\lg\lg n)^{-1/2}$, theorem 1 provides an extension. Although these two theorems are better than the crude results which can be obtained directly from the law of the iterated logarithm, the domain of their applicability is essentially the same as that of the latter, since we use the estimate (4.9) which leads to it.

The following examples, due to Dr. Erdös, show that in general (4.8) is neither necessary nor sufficient for (4.1) or (4.2) even under (4.4) and (4.5).

Example 1. $X_k \equiv 0$ if $k \neq 2^n$; $X_{2^n} = \pm 2^n(\lg\lg n)^{-1}$ each with probability $\frac{1}{2}$.

Example 2. $X_n = 0$ with probability $1 - 2n^{-1}(\lg\lg n)^{-2}$; $X_n = \pm n(\lg\lg\lg n)^{-1}$ each with probability $n^{-1}(\lg\lg n)^{-2}$.

5. Concluding remarks

An opinion may be ventured in conclusion. It is quite possible that the strong law of large numbers will be solved by an approach entirely different from that sketched here. It is even possible that it will be solved by a stroke of great cunning, circumventing all the difficulties inherent in the present methods. Or it may be solved as a result of obtaining sharp asymptotic estimates for probabilities of the form $P(|S_n| > n\epsilon)$ in the general case.

It would appear from the necessary and sufficient conditions given in section 3 that such estimates may be indispensable, but this is not true in the special case of identically distributed random variables (see section 2). However, one thing should be said: the appraisal of such probabilities *is* one of the fundamental problems of the theory of probability, and any real progress in this direction will be of more importance than the solution of a specific problem.

"One hates that power does not come from oneself,
but one does not care if the task is done by oneself."

—Confucius

REFERENCES

[1] A. N. Kolmogorov, "Grundbegriffe der Wahrscheinlichkeitsrechnung," *Ergebnisse der Mathematik*, Vol. 2 (1937); English translation, Chelsea, 1950.

[2] M. Fréchet, *Recherches Théoriques Modernes sur la Théorie des Probabilités*, Book 1, Gauthier-Villars, Paris, 1937.

[3] W. Feller, "A limit theorem for random variables with infinite moments," *Amer. Jour. Math.*, Vol. 58 (1946), pp. 257–262.

[4] T. Kawata, "On the strong law of large numbers," *Proc. Imperial Acad. Tokyo*, Vol. 16 (1940), pp. 109–112.

[5] U. V. Prohorov, "On the strong law of large numbers," (in Russian), *Doklady Akad. Nauk.*, Vol. 69 (1949), pp. 607–610.

[6] P. Lévy, *Théorie de l'Addition des Variables Aléatoires*, Gauthier-Villars, Paris, 1937.

[7] A. N. Kolmogorov, "Sur la loi forte de grandes nombres," *C. R. Acad. Sci., Paris*, Vol. 191 (1930), pp. 910–911.

[8] H. D. Brunk, "The strong law of large numbers," *Duke Math. Jour.*, Vol. 15 (1948), pp. 181–195.

[9] W. Feller, "Generalization of a probability limit theorem of Cramér," *Trans. Amer. Math. Soc.*, Vol. 54 (1943), pp. 361–372.

[10] M. Loève, "Remarques sur la convergence presque sure," *C. R. Acad. Sci., Paris*, Vol. 230 (1950), pp. 52–53; see also his Symposium paper in the present volume.

[11] J. L. Doob, "Application of the theory of martingales," *Le Calcul des Probabilités et Ses Applications, Colloques Internationaux du Centre National de la Recherche Scientifique*, No. 13 (1949), pp. 23–27.

[12] A. N. Kolmogorov, "Ueber die Summen durch den Zufall bestimmter unabhängiger Grössen," *Math. Annalen*, Vol. 99 (1928), pp. 300–319; Vol. 102 (1930), pp. 484–488.

[13] W. Feller, "Ueber das Gesetz der grossen Zahlen," *Acta Litt. ac Sci., Sect. Sci. Math., Szeged*, Vol. 8 (1936–37), pp. 191–201.

[14] J. Marcinkiewicz and A. Zygmund, "Sur les fonctions indépendantes," *Fund. Math.*, Vol. 29 (1937), pp. 60–90.

[15] K. L. Chung, "Note on some strong laws of large numbers," *Amer. Jour. Math.*, Vol. 69 (1947), pp. 189–192.

[16] A. N. Kolmogorov, "Ueber das Gesetz des iterierten Logarithms," *Math. Annalen*, Vol. 101 (1929), pp. 126–135.

Reprinted from MEMOIRS OF THE AMERICAN MATHEMATICAL SOCIETY, Number 6.
With permission from the American Mathematical Society.

ON THE DISTRIBUTION OF VALUES OF SUMS OF RANDOM VARIABLES

By

K. L. Chung* and W. H. J. Fuchs

1. Notation and Summary. $X_1, X_2, \ldots$ are independent, identically distributed random vectors in Euclidean space of k dimensions $(k=1,2,3)$. The distribution function of X_1 is $F(x)$ (or $F(x,y)$ or $F(x,y,z)$ as $k=1,2$, or 3), the characteristic function of X_1 is $\phi(u)$ $(\phi(u,v), \phi(u,v,w))$. $S_n = X_1+X_2 \ldots +X_n$. $|Y|$ denotes the maximum of the absolute values of the components of the vector Y.

The value b is possible, if to every $\varepsilon > 0$ there is an n such that $\Pr\{|S_n-b| < \varepsilon\} > 0$.

The value b is recurrent if for every $\varepsilon > 0$
$\Pr\{|S_n-b| < \varepsilon$ for an infinity of $n\} = 1$.

Theorem 1. *Either no value is recurrent or all possible values are recurrent.*

Theorem 2. *There are recurrent values, if and only if for* $h > 0$

$$\sum_{n=1}^{\infty} \Pr\{|S_n| < h\} = \infty \qquad (1.1).$$

Theorem 3. *There are recurrent values, if for some* $\alpha > 0$

$$\overline{\lim_{\rho \to 1-0}} \int_{-\alpha}^{\alpha} \cdots \int \frac{du \ldots}{1-\rho\phi(u,\ldots)} = \infty \, .^{**} \qquad (1.2)$$

(The number of integrations equals the dimension k of the vectors X_n.)

If for some $\alpha > 0$ *and* $0 < \rho < 1$

$$\int_{-\alpha}^{\alpha} \cdots \int \frac{du \ldots}{1-\rho\phi(u,\ldots)} < K < \infty \qquad (1.3)$$

then there are no recurrent values.

Since the real part of the integrand in (1.2) is positive (see § 3), we obtain by an application of Fatou's Lemma the following

---

*In connection with an ONR project.

**The limit of the left hand side exists, but this is not needed here.

Received by the editors February 20, 1950.

Corollary. A sufficient condition for the existence of recurrent values is

$$\int_{-\alpha}^{\alpha}\cdots\int \frac{du\ldots}{1-\phi(u,\ldots)} = \infty$$

for some $\alpha > 0$.

Theorem 4. If $X_1, X_2 \ldots$ are independent, identically distributed real-valued random variables whose distribution function $F(x)$ satisfies

$$\int_{-\infty}^{\infty} |x|\,dF < \infty, \quad \int_{-\infty}^{\infty} x\,dF = 0 \tag{1.4}$$

then every real number is recurrent, unless all values assumed by X_1 are integral multiples of a fixed number. In this case all (possible and) recurrent values are given by $b=n\lambda$ $(n=0, \pm 1, \pm 2 \ldots)$.

In particular, under the hypotheses of Theorem 4

$$\overline{\lim}\; S_n = \infty$$

with probability one. This result is interesting in view of the fact that W. Feller* proved the existence of a distribution satisfying (1.4) and such that for arbitrarily small $\eta > 0$

$$\Pr\{S_n < -n\,(\log n)^{-\eta}\} \longrightarrow 1$$

as $n \longrightarrow \infty$.

Theorem 5. If $X_1, X_2 \ldots$ are independent, identically distributed vectors in two-dimensional Euclidean space whose distribution function $F(x,y)$ satisfies

$$\iint_{-\infty}^{\infty} x\,dF = \iint_{-\infty}^{\infty} y\,dF = 0$$

$$\iint_{-\infty}^{\infty} (x^2+y^2)\,dF < \infty, \tag{1.5}$$

*Note on the law of large numbers and 'fair' games. Ann. Math. Statistics 16(1945)PP301-304.

then every possible value is recurrent.

Theorem 6. If $X_1, X_2 \ldots$ are independent, identically distributed random vectors with a genuinely three-dimensional distribution, then no value is recurrent*.

Acknowledgement. The simple proofs of Theorems 1 and 2 are due to Professor W. Feller. We are also indebted to him for much valuable advice on other points.

2. Proof of Theorem 1. α) Obviously every recurrent value is also possible.

β) If b is a recurrent value and c is a possible value, then $b-c$ is recurrent: Suppose the contrary. Then for some $\varepsilon > 0$

$q = \Pr\{|S_n-(b-c)| < 2\varepsilon$ for a finite number of indices n only$\} > 0$. Since c is possible there is an index k such that $p = \Pr\{|S_k - c| < \varepsilon\} > 0$. Then

$$\begin{aligned} &\Pr\{|S_n - b| < \varepsilon \text{ for a finite number of } n \text{ only}\} \\ &> \Pr\{|S_k - c| < \varepsilon,\ |S_{k+n} - S_k - (b-c)| < 2\varepsilon \text{ for a finite number of } n \text{ only}\} \\ &> p \cdot q > 0, \end{aligned} \tag{2.1}$$

since the distribution of $S_{k+n} - S_k = X_{k+1} + \ldots + X_{k+n}$ is the same as that of S_n and independent of S_k. But (2.1) contradicts the fact that b is recurrent.

γ) Lemma 1. The set T of all recurrent values forms a closed, additive group.

The definition of a recurrent value implies that T is closed. If b and c are recurrent, then by α) and β) $b-c$ is recurrent. This proves the group property.

Corollary. For a one-dimensional distribution there are the following three possibilities:

1. T is the empty set. 2. T is the set of all real numbers. 3. T is the set of all integral multiples of a number λ. For these are the only closed additive groups of real numbers.

δ) If T is not empty, then $0 \in T$, by γ). If c is any possible value, then $0-c=-c \in T$, by β). Hence $c \in T$, by β). This proves Theorem 1.

Proof of Theorem 2. α) If for any $h > 0$

---

*Theorems 5 and 6 generalize results of Pólya, Ueber eine Aufgabe der Wahrscheinlichkeitsrechnung betreffend die Irrfahrt in Straßennetz 84(1921)pp.149-160.

$$\sum_{n=1}^{\infty} \Pr\{|S_n| < h\} < \infty, \tag{2.2}$$

then by the Borel-Cantelli Theorem $\Pr\{|S_n| < h$ for an infinity of $n\} = 0$. In particular 0 is not recurrent and so T is empty.

β) (1.1) implies

$$q(2h) = \Pr\{|S_n| \geq 2h \ (n=1,2,\ldots)\} = 0.$$

Let

$$r(h) = \Pr\{|S_n| < h \text{ for a finite number of } n \text{ only}\}.$$

Then

$$1 \geq r(h) \geq \sum_{k=1}^{\infty} \Pr\{|S_k| < h, \ |S_{k+n}-S_k| \geq 2h, \qquad (n=1,2,\ldots)\}$$

$$= \sum_{k=1}^{\infty} \Pr\{|S_k| < h\} \, q(2h),$$

since $S_{k+n}-S_k$ is independent of S_k and has the same distribution as S_n. This contradicts (1.1) unless $q(2h) = 0$.

γ) Suppose now that (1.1) holds for every $h > 0$. Then

$$r(h) = \sum_{m>1/h} \sum_{k=1}^{\infty} \Pr\{|S_k| < h - \tfrac{1}{m}, \ |S_{k+n}| \geq h (n=1,2,\ldots)\} + q(h)$$

$$\leq \sum_{m>1/h} \sum_{k=1}^{\infty} \Pr\{|S_k| < h - \tfrac{1}{m}\} \Pr\{|S_{k+n}-S_k| \geq \tfrac{1}{m} \ (n=1,2,\ldots)\}$$

$$= \sum_{m>1/h} \sum_{k=1}^{\infty} \Pr\{|S_k| < h - \tfrac{1}{m}\} \, q(\tfrac{1}{m}) = 0.$$

This concludes the proof of Theorem 2.

3. Proof of Theorem 3. We give the proof for the case of a two-dimensional distribution. The proofs for other dimensions differ only in trivial details.

Let the components of S_n be P_n, Q_n.

$$g(s,t) = \Pr\{|P_n| < s, \ |Q_n| < t\}$$

is a non-decreasing function of s and t. Hence

$$h^2 \Pr\{|S_n| < h\} = h^2 g(h,h) \geq \int_0^h ds \int_0^h dt\, g(s,t).$$

By a well-known formula the right hand side is equal to

$$\frac{1}{\pi^2} \iint_{-\infty}^{\infty} \frac{1-\cos hu}{u^2} \frac{1-\cos hv}{v^2} \phi^n(u,v)dudv$$

Now if $0 \leq \rho < 1$,

$$\frac{1}{\pi^2} \iint_{-\infty}^{\infty} \frac{1-\cos hu}{u^2} \frac{1-\cos hv}{v^2} \frac{du\,dv}{1-\rho\phi(u,v)}$$

$$\geq \frac{A^2(h)}{\pi^2} \iint_{-h^{-1}}^{h^{-1}} \frac{du\,dv}{1-\rho\phi(u,v)}$$

since $\dfrac{1-\cos hu}{u^2} \geq A(h) > 0$ for $|u| \leq h^{-1}$. Hence

$$\sum_{n=0}^{\infty} \rho^n \Pr\{|S_n| < h\} \geq \frac{A^2(h)}{\pi^2 h^2} \iint_{-h^{-1}}^{h^{-1}} \frac{du\,dv}{1-\rho\phi(u,v)}.$$

Notice that the integral is real-valued, since $\phi(-u,-v) = \overline{\phi(u,v)}$. Since $|\phi| \leq 1$,

$$R\frac{1}{1-\rho\phi} = \frac{1-\rho R\phi}{|1-\rho\phi|^2} \geq \frac{1-\rho}{|1-\rho\phi|^2} > 0.$$

If (1.2) is true for a certain α, we have for $h < \alpha^{-1}$

$$\lim_{\rho \to 1-0} \sum_{n=0}^{\infty} \rho^n \Pr\{|S_n| < h\} = \infty.$$

But (1.2) holds a fortiori if we increase α, hence the above is true for all h. Hence (1.1) is true and there are recurrent values by Theorem 2.

Now let $G_n(x,y)$ be the distribution function of S_n. Then

$$\Pr\{|S_n| < h\} = \iint_{-h}^{h} dG_n(x,y) \le \frac{1}{A^2(h^{-1})} \iint_{-h}^{h} \frac{1-\cos h^{-1} x}{x^2} \frac{1-\cos h^{-1} y}{y^2} dG_n(x,y)$$

$$\le \frac{1}{A^2(h^{-1})} \iint_{-\infty}^{\infty} \frac{1-\cos h^{-1} x}{x^2} \frac{1-\cos h^{-1} y}{y^2} dG_n(x,y)$$

$$= \frac{1}{4A^2(h^{-1})} \int_0^{h^{-1}} ds \int_0^{h^{-1}} dt \int_{-s}^{s} du \int_{-t}^{t} dv\, \phi^n(u,v).$$

Hence

$$\sum_{n=0}^{\infty} \rho^n \Pr\{|S_n| < h\} \le \frac{1}{4A^2(h^{-1})} \int_0^{h^{-1}} ds \int_0^{h^{-1}} dt \int_{-s}^{s} \int_{-t}^{t} \frac{du\,dv}{1-\rho\phi(u,v)}$$

$$\le \frac{h^{-2}}{4A^2(h^{-1})} \iint_{-h^{-1}}^{h^{-1}} \frac{du\,dv}{1-\rho\phi(u,v)}.$$

If (1.3) is true for a certain α, we have for $h^{-1} < \alpha$

$$\lim_{\rho \to 1-0} \sum_{n=0}^{\infty} \rho^n \Pr\{|S_n| < h\} < \infty.$$

But (1.3) holds a fortiori if we decrease α, hence the above is true for all h.

4. Theorem 4 is a consequence of the Corollary of Lemma 1 and of the slightly more general

Theorem 4a. If $X_1, X_2, \ldots$ are independent, identically distributed random variables whose distribution function $F(x)$ satisfies

$$\int_{-Y}^{Y} x \, dF(x) = o(1) \tag{4.1}$$

$$Y \int_{|x|>Y} dF(x) = o(1) \tag{4.2}$$

as $Y \longrightarrow \infty$, then recurrent values exist.

Proof. We notice first that

$$\int_{-Y}^{Y} x^2 dF(x) = o(Y) \tag{4.3}$$

as $Y \longrightarrow \infty$. For writing $F(x) - F(-x) = F^*(x)$,

$$\int_{-Y}^{Y} = \int_{0}^{Y} x^2 dF^*(x) = Y^2(1-F^*(Y))+2\int_{0}^{Y} x(1-F^*(x))dx$$

$$= o(Y)+2\int_{0}^{Y} o(1)dx = o(Y).$$

Now

$$1-\rho \le R(1-\rho\phi(u)) = 1-\rho+\rho . R(1-\phi(u))$$

$$\le 1-\rho+\int_{-\infty}^{\infty} (1-\cos x u) \, dF(x)$$

$$\le 1-\rho+\frac{1}{2}\int_{-1/|u|}^{1/|u|} (xu)^2 \, dF(x) + 2\int_{|x|>1/|u|} dF(x)$$

since $0 \le 1 - \cos x u \le \min(2, \frac{1}{2}(xu)^2)$. Hence, by (4.2) and (4.3)

$$R(1-\rho\phi) \le 1-\rho+ o(u^2|u|^{-1})+o(|u|) = 1-\rho + o(|u|)$$

as $u \longrightarrow 0$. Also

$$|I\rho\phi(u)| \leq \left|\int_{-\infty}^{\infty} \sin x\, u dF\right| \leq \left|\int_{|x| \leq 1/|u|} xu\, dF(x)\right| + 0\left(\int_{|x| \leq 1/|u|} x^2u^2 dF\right) + \int_{|x| > 1/|u|} dF ,$$

since $\sin x u = x u + 0\left((xu)^2\right)$. Hence, using (4.1) and (4.3) $|I\rho\phi(u)| \leq o(|u|)$. Given $\varepsilon > 0$ and $\alpha > 0$ there is an $\alpha_o \leq \alpha$ such that in $0 < u < \alpha_o$

$$R\frac{1}{1-\rho\phi(u)} = \frac{R(1-\rho\phi(u))}{(R(1-\rho\phi(u))^2 + (I\rho\phi(u))^2}$$

$$\geq \frac{1-\rho}{(1-\rho+\varepsilon u)^2+(\varepsilon u)^2} \geq \frac{1-\rho}{4(1-\rho)^2+4(\varepsilon u)^2}$$

and so, using (3.2)

$$\int_{-\alpha}^{\alpha}\frac{du}{1-\rho\phi(u)} = 2R\int_0^{\alpha} \geq 2R\int_0^{\alpha_o} \geq \frac{1}{2}\int_0^{\alpha_o} \frac{1-\rho}{(1-\rho)^2+(\varepsilon u)^2}\, du,$$

as $\rho \longrightarrow 1$ the last integral tends to $\frac{1}{2}\int_0^{\infty} \frac{dv}{1+(\varepsilon v)^2} = \frac{\pi}{4\varepsilon}$.

Since ε is arbitrarily small, (1.2) must hold and the theorem is proved.

Remark. The conditions of Theorem 4 are not necessary and can be varied in several ways. E.g. we can replace (4.2) by

$$Y\int_{|x| > Y} dF(x) = O(1),$$

if at the same time (4.1) is strengthened to

$$\int_{-Y}^{Y} dF(x) = 0,$$

i.e. the distribution of X_1 is symmetrical. But some condition like (4.2) is necessary,

even for symmetrical distributions, as the following example shows. Let $0 < c < 1$,

$$F'(x) = F'_c(x) \begin{cases} = 0 & (|x| < 1) \\ = \frac{1}{2} c|x|^{-c-1} & (|x| > 1) \end{cases} \tag{4.4}$$

Then, for $u > 0$,

$$\begin{aligned}
\phi(-u) = \phi(u) &= c \int_1^\infty x^{-c-1} \cos xu \, dx \\
&= 1 - c \int_1^\infty (1-\cos xu) x^{-c-1} dx \\
&= 1 - cu^c \int_u^\infty (1-\cos t) t^{-c-1} dt \\
&= 1 - cu^c \left(\int_0^\infty - \int_0^u \right) \\
&\leq 1 - Au^c.
\end{aligned}$$

Hence

$$\int_{-\alpha}^{\alpha} \frac{du}{1-\rho\phi(u)} \leq 2 \int_0^{\alpha} \frac{du}{Au^c} = K < \infty,$$

so that no value is recurrent. Here

$$Y \int_{|x| > Y} dF = kY^{1-c} .$$

5. Proof of Theorem 5. Under the hypothesis

$$1-\phi(u,v) = \iint_{-\infty}^{\infty} (1-e^{i(xu+yv)})dF(x,y)$$

$$= O\left(\iint_{-\infty}^{\infty} (xu+yv)^2 dF\right).$$

Hence

$$|1-\phi(u,v)| \leq K \iint_{-\infty}^{\infty} (u^2+v^2)(x^2+y^2)dF = B(u^2+v^2).$$

$$\iint_{-\alpha}^{\alpha} \frac{dudv}{1-\phi(u,v)} \geq \frac{1}{B} \iint_{-\alpha}^{\alpha} \frac{dudv}{u^2+v^2} = \infty$$

and Theorem 5 follows from the Corollary of Theorem 3.

Remark. The condition (1.5) is not necessary, but it can not be relaxed very much:

Let

$$F(x,y) = \int_{-\infty}^{x}\int^{y} F'_c(\xi)F'_c(\eta)d\xi\, d\eta ,$$

where F'_c is defined by (4.4), but now with $1 < c \leq 2$. A calculation very similar to that in § 4 shows that near the origin

$$\phi(u,v) < (1-k|u|^c)(1-k|v|^c) \qquad (1 < c < 2)$$

$$< 1-K(|u|^c + |v|^c)$$

$$\phi(u,v) = (1+u^2 \log |u| + O(u^2))(1+v^2\log|v| + O(v^2)) \qquad (c=2)$$

Hence, for $c=2$,

$$\iint_{-\alpha}^{\alpha} \frac{du\,dv}{1-\phi(u,v)} = 8 \int_0^{\alpha} du \int_0^{u} \frac{dv}{1-\phi(u,v)}$$

$$\geq A \int_0^{\alpha} du \int_0^{u} \frac{dv}{v^2 \log(1/v)} = A \int_0^{\alpha} \frac{du}{u \log(1/u)} = \infty$$

and there are recurrent values although

$$\iint (x^2+y^2)\, dF(x,y) = \infty.$$

If $1 < c < 2$, then

$$\iint_{-\alpha}^{\alpha} \frac{du\, dv}{1-\rho\phi(u,v)} \leq A \iint_{-\alpha}^{\alpha} \frac{du\, dv}{|u|^c+|v|^c} < 8A \int_0^{\alpha} du \int_0^{u} \frac{dv}{u^c} < \infty$$

and no value is recurrent. But

$$\iint (|x|^b+|y|^b) dF(x,y) = 2 \int_{-\infty}^{\infty} |x|^b F'_c(x) dx < \infty,$$

if $b < c$. Hence the condition (1.5) can not be replaced by

$$\iint (|x|^b+|y|^b) dF(x,y) < \infty$$

with any $b < 2$.

6. Proof of Theorem 6. Our assumption is that there is no plane through the origin such that X_1 lies with probability one in this plane. (The distribution is genuinely three-dimensional). Hence there is a sufficiently large R so that

$$Q = \iiint_{-R}^{R} (ux+vy+wz)^2\, dF(x,y,z) > 0$$

for <u>all</u> u,v,w, with $u^2+v^2+w^2 \neq 0$. For the left hand side can be equal to 0 only if all possible values of X_1 with $|X_1| < R$ lie on the plane $ux+vy+wz = 0$ with probability one. Therefore Q is a positive definite form in u,v,w and hence

$$Q > M(u^2+v^2+w^2).$$

Choose $|u|, |v|, |w| < 1/R$. Then

$$R(1-\phi(u,v,w)) = 2 \iiint_{-\infty}^{\infty} \sin^2 \tfrac{1}{2} (ux+vy+wz)\, dF$$

$$\geq 2 \iiint_{-R}^{R} \geq 2\pi^{-2} Q > A(u^2+v^2+w^2),$$

since in $\max(|x|,|y|,|z|) < R$ $|ux+vy+wz| < 3$ and so

$$|\sin \tfrac{1}{2}(ux+vy+wz)| > \frac{2}{\pi} \cdot \frac{1}{2} |ux+vy+wz|.$$

Hence for $\alpha < 1/R$

$$\iiint_{-\alpha}^{\alpha} \frac{du\,dv\,dw}{1-\rho\phi(u,v,w)} \leq \iiint_{-\alpha}^{\alpha} \frac{du\,dv\,dw}{R(1-\phi)} \leq A \iiint_{-\alpha}^{\alpha} \frac{du\,dv\,dw}{u^2+v^2+w^2} < \infty.$$

Theorem 6 is now a consequence of Theorem 3.

Cornell University

PROBABILITY LIMIT THEOREMS ASSUMING ONLY THE FIRST MOMENT I

By

K. L. CHUNG and P. ERDÖS

In this paper we consider sums of mutually independent, identically distributed random variables. An essential feature is that we assume only that the first moment is zero, or that both its positive and negative parts diverge. Part I here deals with lattice distributions. Perhaps the main results are Theorem 3.1 and Theorem 8. We hope to take up other cases later.

1. Let X be a random variable which assumes only integer values

$$P(X = k) = p_k$$

$$p_k \geq 0 \quad \sum p_k = 1^* \; .$$

A number is said to be a 'possible' value of an integer-valued random variable if its probability is positive. The possible values of X will be denoted by $u_i, i=1,2,\ldots;$ they may be finite or infinite in number. As usual $S_n = \sum_{k=1}^{n} X_k$ where the X_k are mutually independent, each having the same distribution as X.

To avoid minor complications, we shall assume that every integer c is a possible values of S_n if n is sufficiently large: $n \geq n_0(c)$. A set of necessary and sufficient conditions for this is the following:

(1) The u_i are not all of the same sign;

(2) The greatest common divisor of the set of differences $u_i - u_j, i,j=1,2,\ldots$ is equal to 1.

We shall call the following two sets of assumptions (0) and (∞) respectively:

$$(0) \qquad E(|X|) = \sum |k| p_k < \infty, \quad E(X) = \sum k p_k = 0$$

$$(\infty) \qquad \frac{1}{2} E(|X|+X) = \sum_{k=0}^{\infty} k p_k = \infty, \; \frac{1}{2} E(|X|-X) = -\sum_{k=-\infty}^{0} k p_k = \infty .$$

Thus under (0) or (∞) (1) is always satisfied except in the trivial case $X \equiv 0$, which we exclude. If (2) is not satisfied, there are two possibilities: either all possible

---

Received by the editors January 5, 1950.

*In an unspecified summation the index runs from $-\infty$ to $+\infty$.

values of S_n are multiples of an integer > 1; or there exists an integer $m > 1$ and a complete residue class mod. m, $r_1,\dots,r_m$ such that for a fixed j, all possible values of S_{nm+j}, $n=1,2,\cdots$, belong to the same residue class $r_j \pmod m$. It is not difficult to see how our statements and proofs should be modified for these cases.

In the following the letters a, a' denote arbitrary integers, A, A', B positive constants; ε, ε' arbitrarily small constants.

2. In this section we give some simple theorems on the bounds of $P(S_n=a)$. It is well known that under more restrictive assumptions more precise results can be obtained (see Gnedenko [1], van Kampen and Wintner [2], Esseen [3]).

THEOREM 1. Under no assumptions about moments whatsoever,

$$P(S_n=a) \leq An^{-1/2} \tag{1}$$

where A does not depend on n or a. If $E(X^2) = \infty$, then

$$\lim_{n\to\infty} n^{1/2}P(S_n=a) = 0. \tag{2}$$

Proof. The c.f.* of the d.f. † of X is

$$f(x) = \sum p_k e^{ikx}.$$

The c.f. of S_n if $(f(x))^n$, and we have

$$P(S_n = a) = \frac{1}{2\pi}\int_{-\pi}^{\pi} (f(x))^n e^{-iax}dx.$$

Suppose first that n is even: $n=2m$. We have

$$\left|\int_{-\pi}^{\pi} (f(x))^{2m}e^{-iax}dx\right| \leq \int_{-\pi}^{\pi} (|f(x)|^2)^m dx.$$

Now $|f(x)|^2$ is the c.f. of a symmetrical d.f., namely that of $X + X'$ where X,X' are mutually independent and X' has the same distribution as $-X$. Hence we may write

$$|f(x)|^2 = \sum_{k=0}^{\infty} r_k \cos kx, \left(r_k \geq 0,\ \sum_{k=0}^{\infty} r_k = 1\right)$$

*Characteristic function or Fourier-Stieltjes transform.

† Distribution function.

$$= 1 - 2 \sum_{k=0}^{\infty} r_k \sin^2 \frac{kx}{2}.$$

Suppose $r_\ell > 0$. If $x \leq \pi \ell^{-1}$,

$$2 \sum_{k=0}^{\infty} r_k \sin^2 \frac{kx}{2} \geq 2\left(\sum_{k=0}^{\ell} r_k \left(\frac{kx}{\pi}\right)^2\right) \geq \frac{2x^2}{\pi^2} \sum_{k=0}^{\ell} r_k k^2 > Ax^2$$

$$|f(x)|^2 < 1 - A x^2.$$

Hence

$$\int_{-\pi}^{\pi} |f(x)|^{2m} dx < \int_{-\pi\ell^{-1}}^{\pi\ell^{-1}} (1-A x^2)^m dx + \int_{\pi\ell^{-1}<|x| \leq \pi} |f(x)|^{2m} dx.$$

It is known that if $\eta < |x| \leq \pi$, then $|f(x)| < 1 - \varepsilon(\eta)$. Therefore we have

$$\int_{-\pi}^{\pi} |f(x)|^{2m} dx \leq \int_{-\pi\ell^{-1}}^{\pi\ell^{-1}} e^{-A mx^2} dx + O((1-\varepsilon)^m).$$

(1) follows for even n. Noticing that $|f(x)|^n \leq |f(x)|^{n-1}$ we see that the same proof goes through for an odd n.

To prove (2), notice that the assumption $E(X^2) = \infty$ implies that $E((X+X')^2) = \infty$. Hence

$$\sum_{k=0}^{\infty} k^2 r_k = \infty,$$

and the A' in the foregoing can be taken arbitrarily large. q.e.d.

A lower bound for $P(S_n=a)$, under the assumption (0) or (∞), will be given in Theorem 2.2; we shall also show that our estimate is close to the best possible by exhibiting an example in Theorem 2.3. In one special case, however, we can prove a much stronger result, and this is Theorem 2.1.

THEOREM 2.1. <u>If the</u> d.f. <u>of</u> X <u>is symmetrical</u>, <u>and</u> $E(|X|) < \infty$, <u>then</u>

$$\lim_{n\to\infty} n\, P(S_n=a) = \infty. \tag{3}$$

Proof. Since $p_k=p_{-k}$, $f(x)$ is real. Since $f(0) = 1$ and $f(x)$ is continuous, there exists a $\delta > 0$ such that if $|x| < \delta$, $f(x) > 0$. We have

$$P(S_n=a) = \frac{1}{2\pi} \int_{-\pi}^{\pi} (f(x))^n \cos ax dx,$$

$$\geq \frac{1}{4\pi} \int_{-\delta}^{\delta} (f(x))^n dx - O((1-\epsilon)^n),$$

if $\delta < \frac{\pi}{3a}$. As in the proof of Theorem 1, we can write

$$1-f(x) = \sum_{k=0}^{\infty} 2r_k \sin^2 \frac{kx}{2}.$$

Since $\sum_{k=0}^{\infty} kr_k < \infty$,

$$\lim_{x \to 0} \frac{1}{|x|} \sum_{k=0}^{\infty} r_k \sin^2 \frac{kx}{2} = 0.$$

Hence given $\varepsilon > 0$, if $|x| < \delta_0(\varepsilon) < \delta$, $1 - f(x) \leq \varepsilon|x|$. Now

$$\int_{-\delta}^{\delta} (f(x))^n dx \geq \int_{-\delta_0}^{\delta_0} (1-\varepsilon|x|)^n dx = \frac{2-2(1-\delta_0\varepsilon)^{n+1}}{(n+1)\varepsilon}.$$

Since ε is arbitrary (3) follows.

THEOREM 2.2. Under (0) or (∞) we have for every $\varepsilon > 0$

(4) $\quad P(S_n=a) \geq (1-\epsilon)^n$

if $n \geq n(\varepsilon, a)$.

Proof. If the possible values of X are bounded, then $E(X^2) < \infty$. In this case it is well known and also easy to show that

$$\lim_{n \to \infty} n^{1/2} P(S_n=a) = A < \infty.$$

This is a much sharper result than (4). Hence we may assume that there are possible values of arbitrarily large magnitude.

Given $\varepsilon > 0$, there exist arbitrarily large z_1 and z_2 such that

$$\sum_{-z_1}^{z_2} p_k > 1-\epsilon$$

and if $k > z_2, p_k < \varepsilon$. Now choose h' so large that

$$\left| \sum_{-h'}^{0} kp_k \right| > \sum_{0}^{z_2} kp_k$$

this is possible under (0) or (∞). Also there is a unique h such that

$$\sum_{0}^{h-1} kp_k < \Big| \sum_{-h'}^{0} kp_k \Big| \leq \sum_{0}^{h} kp_k.$$

Then $h > z_2$ and

$$hp_h \geq \sum_{-h'}^{h} kp_k = C \geq 0.$$

Define $p'_k = p_k$ if $k \neq h$, but

$$p'_h = p_h - Ch^{-1} \geq 0.$$

Then

$$b = \sum_{-h'}^{h} p'_k > 1 - \varepsilon - p_h > 1 - 2\varepsilon$$

$$\sum_{-h'}^{h} kp'_k = 0.$$

Now define a random variable X' as follows:

$$P(X' = k) = p'_k b^{-1} \qquad \text{if} \qquad -h' \leq k \leq h$$

$$= 0 \qquad\qquad \text{otherwise.}$$

Let $S'_n = \sum_{k=1}^{n} X'_k$ where the X'_k are mutually independent and each has the same distribution as X'. Since $p'_k \leq p_k$ for all k

$$P(S'_n = a) \geq b^{-n} P(S_n = a \mid -h' \leq X_k \leq h \text{ for } 1 \leq k \leq n).*$$

Hence

$$P(S_n = a) \geq P(S_n = a; -h' \leq X_k \leq h \text{ for } 1 \leq k \leq n)$$

$$\geq P(-h' \leq x \leq h)^n P(S_n = a \mid -h' \leq X_k \leq h \text{ for } 1 \leq k \leq n)$$

$$\geq (1-\varepsilon)^n b^n P(S'_n = a) \geq (1-\varepsilon)^n (1-2\varepsilon)^n A n^{-1/2}$$

where A depends on ε by definition of X'. This being true for all ε is equivalent to (4).

The idea of truncation in the preceding proof is due to Shizuo Kakutani. Theorem 2.2 was first proved under (0) by W. H. J. Fuchs using a result in Chung and Fuchs [4], namely

---

*$P(E|F)$ denotes the conditional probability of E under the hypothesis F.

$$\sum_{n=1}^{\infty} P(S_n = a) = \infty.^* \tag{5}$$

A similar proof using (5) was also given by Kakutani. We sketch the latter proof as follows.

From (5) it follows by the Cauchy-Hadamard criterion

$$\overline{\lim_{n\to\infty}} \quad (P(S_n=0))^{1/n} = 1.$$

Hence given $\varepsilon > 0$, there exists arbitrarily large m such that

$$P(S_m = 0) \geq (1-\varepsilon)^m.$$

Consequently for all integers $k > 0$,

$$P(S_{km} = 0) \geq (1-\varepsilon)^{km}.$$

We can also choose the aforesaid m so large that

$$\min_{m<v\leq 2m} P(S_v=a) = A' > 0.$$

Now fix m. If $n=(k+1)m+r$, $k > 0$, $\varepsilon < r < m$, we have

$$P(S_n=a) \geq P(S_{m+r} = a)P(S_{km} = 0)$$

$$\geq A'(1-\varepsilon)^{km} \geq A'(1-\varepsilon)^{-m}(1-\varepsilon)^n. \qquad \text{q.e.d.}$$

THEOREM 2.3. <u>We can construct an example satisfying</u> (0) <u>and such that for every given</u> $B > 0$ <u>there exists a sequence</u> $\{n_v\}$ <u>for which</u>

$$P(S_{n_v} \geq 0) = O(n_v^{-B}).$$

Proof. Let $A_v, v=1,2,\dots$ be a sequence of positive integers increasing to ∞ so fast that for every $\varepsilon > 0$,

$$A_v = O(A_{v+1}^{\varepsilon}).$$

Define

$$X = \begin{cases} -1 \text{ with prob. } \frac{1}{2} \\ \\ A_v \text{ with prob. } 2^{-v}A_v^{-1} \end{cases} \qquad v=2,3,\cdots .$$

Then $E(X) = 0$. If k is sufficiently large

$$P(\max_{1\leq v\leq n} X_v > A_k) \leq n \sum_{v=k+1}^{\infty} \frac{1}{2^v A_v} \leq \frac{n}{A_{k+1}}.$$

---

*However, the assumption (∞) does not imply the truth of (5) (see [4]); thus the following proof does not hold under (∞).

Let

$$X^* = \begin{cases} X \text{ if } X \leq A_k \\ 0 \text{ if } X > A_k \end{cases}.$$

$S^*_n = \sum_{v=1}^{n} X^*_v$ where the X^*_v are mutually independent and each has the distribution of X^*.

We have

$$E(X^*) = -2^{-(k+1)},\ E(X^{*2}) = O(A_k)$$

$$P(S_n \geq 0) \mid \max_{1\leq v\leq n} X_v \leq A_k) = P(S^*_n \geq 0)$$

$$\leq P(|S^*_n - E(S^*_n)| \geq |E(S^*_n)|).$$

Let m be an integer > 0, a routine computation shows that

$$E(|S^*_n - E(S^*_n)|^{2m} \leq K_m(\sum_{v=1}^{n} E(X^* + 2^{-(k+1)})^2)^m$$

$$\leq K'_m A_k^m n^m$$

where K_m, K'_m are two positive constants depending only on m. Hence

$$P(|S^*_n - E(S^*_n)| \geq |E(S^*_n)|) \leq O(A_k^m 2^{(k+1)2m} n^{-m}).$$

Now choose

$$n_k \sim A_{k+1}^{\varepsilon}$$

we have, by the property of the sequence A_k,

$$P(S_{n_k} \geq 0) \leq P(\max_{1\leq v\leq n_k} X_v > A_k) + P(S_{n_k} \geq 0 | \max_{1\leq v\leq n_k} X_v \leq A_k)$$

$$\leq O(n_k^{-1/\varepsilon+1} + n_k^{-m+\varepsilon}) = O(n_k^{-B})$$

by choice of ε and m.

Theorem 2.3 should be compared with a result due to Feller [5].

3. The theorems in § 2 were proved by fairly standard analytical methods. We are unable to prove the theorems in this section by similar methods, except in the case where the d.f. of X is symmetrical, i.e., the c.f. is a real-valued function. In this case we have as before

$$|P(S_n=a) - P(S_n=a')| \le \int_{-\delta}^{\delta} |\cos ax - \cos a'x|\ |f(x)|^n\, dx + O((1-c)^n).$$

Choosing δ so small that $\cos ax > 0$, $f(x) > 0$, and $|\cos ax - \cos a'x| < \varepsilon' \cos ax$ for $|x| \le \delta$, we have

$$|P(S_n=a) - P(S_n=a')| \ge \epsilon' \int_{-\delta}^{\delta} \cos ax (f(x))^n\, dx + O((1-c)^n).$$

On account of Theorem 2.2 it follows that

$$P(S_n=a) - P(S_n=a') = o\,(P(S_n=a))$$

which is equivalent to Theorem 3.1 below. We have not been able to prove the theorem by this method when $f(x)$ is not real-valued. Another relevant remark is the following: if instead of the individual probabilities $P(S_n=a)$ we consider their sums, then it follows from a theorem due to Doeblin [11] on Markov chains that

$$\lim_{n\to\infty} \frac{\sum_{k=1}^{n} P(S_k=a)}{\sum_{k=1}^{n} P(S_k=a')} = 1.$$

THEOREM 3.1. Under (0) or (∞)

$$\lim_{n\to\infty} \frac{P(S_n=a)}{P(S_n=a')} = 1.$$

Proof. For some $k, u_1-u_0, \cdots, u_k-u_0$ have g.c.d. 1. Thus there exist integers c_i' and c_i such that

$$a'-a = \sum_{i=1}^{k} c_i'(u_i-u_0) = \sum_{i=0}^{k} c_i u_i, \quad \sum_{i=0}^{k} c_i = 0.$$

Let $P(X=u_i) = q_i$. Corresponding to every representation of a in the form

(1) $$a = \sum_{i=0}^{\ell} n_i u_i, \qquad n_i \ge 0, \qquad \sum_{i=0}^{\ell} n_i = n$$

there is a realization of the value a by $X_1+\cdots+X_n$ with probability

(2) $$\frac{n!}{n_0!\cdots n_\ell!} \prod_{i=0}^{\ell} q_i^{\,n_i}$$

when n_i of the X's assume the value u_i. The total probability of a is thus

$$\sum \frac{n!}{n_0!\cdots n_\ell!} \prod_{i=0}^{\ell} q_i^{n_i}$$

where the sum runs over all representations (1). Now write this sum as

(3) $$\sum = \sum_1 + \sum_2$$

where in $\sum_1$ the conditions

$$|n_i - nq_i| < \varepsilon n, \qquad 0 \leq i \leq \ell$$

are satisfied, while $\sum_2$ is the rest.

Consider the event $X = u_i$ with probability q_i; n_i is the number of its occurrences in n mutually independent, identical trials. It is well known that the probability that $|n_i - nq_i| > \varepsilon n$ is

$$O(e^{-\varepsilon'^2 n}).$$

Hence

(4) $$\sum_2 \leq (\ell+1)O(e^{-\varepsilon'^2 n}) = o(P(S_n = a))$$

by Theorem 2.2, for every $\varepsilon > 0$.

Now consider a representation (1) with $|n_i - nq_i| < \varepsilon n$ for $0 \leq i \leq \ell$. If ε is sufficiently small and n sufficiently large, we have $n_i > n(q_i - \varepsilon) > \varepsilon n > |c_i|$. Corresponding to every representation of a in the form (1) there is a representation of a' in the form

(5) $$a' = \sum_{i=0}^{k} (n_i + c_i)u_i + \sum_{i=k+1}^{\ell} n_i u_i = \sum_{i=0}^{\ell} n_i' u_i$$

where $|n_i' - nq_i| < 2\varepsilon n$. The ratio of two such corresponding probabilities is

$$= \frac{n_0!\cdots n_k!}{n_0'!\cdots n_k'!} q_0^{n_0' - n_0} \cdots q_k^{n_k' - n_k}.$$

If $m' > m$, $|m - nq| < \varepsilon n$, $|m' - nq| < 2\varepsilon n$, we have

$$\frac{m!}{m'!} q^{m'-m} \frac{(qn)(qn)\cdots(qn)}{m'(m'-1)\cdots(m+1)} n^{m-m'}$$

$$\leq \left(\frac{qn}{(q-2\varepsilon)n}\right)^{m'-m} n^{m-m'} = \left(\frac{1}{1-2\varepsilon}\right)^{m'-m} n^{m-m'}$$

$$\frac{m!}{m'!} q^{m'-m} \geq \left(\frac{1}{1+2\varepsilon}\right)^{m'-m} n^{m-m'}.$$

Since $\sum_{i=0}^{\ell} n_i' = \sum_{i=0}^{\ell} n_i$ it follows that

$$(1-\varepsilon')^C \leq \lambda \leq (1+\varepsilon')^C$$

where $C = \sum_{i=0}^{k} |c_i|$. Since ε' is arbitrarily small we have $\lim_{n\to\infty} \lambda = 1$.

Let us write the corresponding formulas (1) and (2) for a':

$$a' = \sum_{i=0}^{\ell} n_i' u_i, \qquad \sum_{i=0}^{\ell} n_i' = 0$$

(6) $$\sum{}' = \sum{}_1' + \sum{}_2'$$

where in $\sum_1'$ the condition $|n_i' - nq_i| < 2\varepsilon n$ are satisfied find $0 \leq i \leq \ell$. We have just proved that

$$\overline{\lim_{n\to\infty}} \frac{\sum_1}{\sum_1'} \leq 1.$$

Using (4) we conclude that

$$\overline{\lim_{n\to\infty}} \frac{P(S_n=a)}{P(S_n=a')} = 1.$$

Since a and a' are interchangeable we obtain Theorem 3.

THEOREM 3.2. <u>For those values of</u> n <u>for which</u>

(7) $$P(S_n=a) \geq n^{-B}$$

<u>for some fixed</u> $B > 0$, <u>we have for every</u> $\varepsilon > 0$

(8) $$|P(S_n=a) - P(S_n=a')| \leq P(S_n=a)An^{-1/2+\varepsilon}$$

<u>where</u> A <u>may depend on</u> a, a' <u>but not on</u> n.

Proof. In (3) we re-define $\sum_1$ to be the sum of those terms for which

$$|n_i - nq_i| < n^{1/2+\varepsilon}, \qquad 0 \leq i \leq \ell.$$

As before, let (5) correspond to (1), but now we assume n so large that $n^{1/2+\varepsilon} > |C_1|$, so that

$$|n_i' - nq_i| < 2n^{1/2+\varepsilon}.$$

We re-define $\sum_1'$ in (6) to be the sum of those terms for which this is true. By well known estimates on the binomial distribution we have

$$\sum_2 = O(e^{-An^{\varepsilon}}), \quad \sum_2' = O(e^{-An^{\varepsilon}}). \tag{9}$$

Now consider the difference of two corresponding probabilities (1) and (6)

$$d = \frac{n!}{n_0! \cdots n_\ell!} q_0^{n_0} \cdots q_\ell^{n_\ell} (1-\lambda).$$

If $m = nq+r$, $m' = nq+r'$ where $|r-r'| \le C$ and $|r| \le n^{1/2+\varepsilon}$, $|r'| \le 2n^{1/2+\varepsilon}$, an easy application of Stirling's formula yields

$$\frac{m!}{m'!} q^{m'-m} = n^{r-r'}(1+O(n^{-1/2+3\varepsilon})), \qquad \lambda = 1+O(kn^{-1/2+3\varepsilon}).$$

Since $\sum_{i=0}^{k} (r_i - r_i') = 0$ we have

$$|d| \le \frac{n!}{n_0! \cdots n_\ell!} q_0^{n_0} \cdots q_\ell^{n_\ell} O(n^{-1/2+3\varepsilon}).$$

Hence

$$|P(S_n = a) - P(S_n = a')| \le O(\sum_1 . n^{-1/2+3\varepsilon}) + \sum_2 + \sum_2'.$$

The first term on the right is $O(P(S_n = a).n^{-1/2+3\varepsilon})$, and the other two terms by (7) and (9) are of smaller order of magnitude. Thus (8) follows.

THEOREM 4. Under (0) or (∞)

$$\lim_{n \to \infty} \frac{P(S_n = a)}{P(S_{n+1} = a')} = 1.$$

Proof. For every representation of a in the form (1), there is a representation of $a+u_0$ in the following form

$$a+u_0=(n_0+1)u_0+\sum_{i=1}^{\ell} n_i u_i .$$

The corresponding probability is

$$\frac{(n+1)!}{(n_0+1)!n_1!\cdots n_\ell!}\, q_0^{n_0+1}\, q_1^{n_1}\cdots q_\ell^{n_\ell} .$$

The ratio of this to (2) is $(n_0+1)/(n+1)q_0$. If $|n_0-nq_0|<\varepsilon n$, this ratio is between $1-\varepsilon/q$ and $1+\varepsilon/q_0$ as $n\longrightarrow\infty$. The range at values of n_0 such that $|n_0-nq_0|>\varepsilon n$ can be neglected as before. It follows exactly as in the proof of Theorem 3 that

$$\overline{\lim_{n\longrightarrow\infty}}\ \frac{P(S_n=a)}{P(S_{n+1}=a+u_0)}\leq 1.$$

By virtue of Theorem 3 this gives

$$\overline{\lim_{n\longrightarrow\infty}}\ \frac{P(S_n=a)}{P(S_{n+1}=a)}\leq 1.$$

Considering $a-u_0$ instead of $a+u_0$ in the above in a similar manner we arrive at

$$\overline{\lim_{n\longrightarrow\infty}}\ \frac{P(S_n=a)}{P(S_{n-1}=a)}\leq 1.$$

These last two relations combined are equivalent to Theorem 4.

We remark that Theorem 4 can be proved in the same way as sketched above for Theorem 3.1, when $f(x)$ is real-valued. It would also seem that we might be able to deduce Theorem 4 directly from Theorem 3.1, but a trivial argument gives only the following. Since

$$P(S_{n+1}=a)=\sum_{a'=-\infty}^{\infty} P(S_n=a')\,P(X=a-a')$$

$$\geq \sum_{a'=-A}^{A} P(S_n=a')P(X=a-a').$$

It follows easily, using Theorem 3.1, that

$$\underline{\lim_{n\longrightarrow\infty}}\ \frac{P(S_{n+1}=a)}{P(S_n=a)}\geq 1.$$

But the other half of the result seems difficult.

4. In this section we study the number of a-values in the sequence $S_1, \ldots, S_n$. A very special case has been treated more or less completely by Chung and Hunt [6]. More general cases, in which the existence of certain moments are assumed, have been considered by Feller [7] and Chung [8].* In this paper we are considering a more general situation and precise results are not hoped for at this moment. However, we shall prove the relevant Theorem 8 whose truth would perhaps be considered evident but whose proof, as far as we can make it, is by no means simple. Theorem 7 gives the true bounds within an ε power.

Define

$$Y_k = \begin{cases} 1 \text{ if } S_k = a \\ 0 \text{ if } S_k \neq a \end{cases}$$

$$E(Y_k) = P(S_k = a) = m_k$$

$$T_n = \sum_{k=1}^{n} Y_k$$

$$E(T_n) = M_n = \sum_{k=1}^{n} m_k$$

and similarly Y'_k, m'_k, T'_n, M'_n, for a'.

THEOREM 5. <u>Under</u> (0), <u>for every</u> $\varepsilon > 0$,

$$P(|T_n - T'_n| > M_n^{3/4+\varepsilon} \text{ i.o.}^{**}) = 0.$$

Proof. By Theorem 3.1 and the fact that $M_n \longrightarrow \infty$ as $n \longrightarrow \infty$

$$E(|T_n - T'_n|^2) = E\Big(\sum Y_k^2 + \sum Y'^2_k + \sum_{j \neq k} Y_j(Y_k - Y'_k) + \sum_{j \neq k} Y'_j(Y'_k - Y_k)\Big) \dagger$$

$$\ll \sum m_j + \sum m'_j + \sum m_j \sum_{k \neq j} |m_{j-k} - m'_{j-k}| + \sum m'_j \sum_{k \neq j} |m'_{j-k} - m_{j-k}| \ddagger$$

$$\ll M_n + M_n \sum |m_k - m'_k|.$$

---

*The results in [8] are stated for the number of crossings of the values a, but in the case of an integer-valued random variable they can be easily translated into the number of a-values. ** i.o. stands for 'infinitely often' or 'for infinitely many values of the index.'

† Henceforth in an unspecified summation the index runs from 1 to n. ‡ $u_n \ll v_n$ means $u_n = O(v_n)$

According to Theorem 3.2, the m_k in the last sum can be divided into two classes: either $m_k \le k^{-2}$, the sum over such k being $O(1)$, or the estimate (8) holds. Hence

$$E(|T_n-T'_n|^2) \le O(M_n) + O(M_n \sum m_k k^{-(1-\epsilon)/2}).$$

By Hölder's inequality

$$\sum m_k k^{-(1-\epsilon)/2} \le (\sum m_k^{2/(1+2\epsilon)})^{1/2+\epsilon} (\sum (k^{-(1-\epsilon)/2})^{2/(1-2\epsilon)})^{1/2-\epsilon}$$

$$\le A(\sum m_k^{2/(1+2\epsilon)})^{1/2+\epsilon} \le AM_n^{1/2+\epsilon}.$$

By Chebychev inequality

$$(10) \qquad P(|T_n-T'_n| > M_n^{3/4+\epsilon}) \le M_n^{-\epsilon}.$$

Since $m_n \longrightarrow 0$ by Theorem 1, we can choose an increasing sequence n_k such that

$$M_{n_k} \sim k^{(1+\epsilon)/\epsilon}.$$

Now suppose that for some $n, n_k < n \le n_{k+1}$ we have

$$(11) \qquad |T_n-T'_n| > 2M_n^{3/4+\epsilon}.$$

Let n be the smallest such integer, for which (11) is true, then either Y_n or Y'_n must be 1, hence $S_n=a$ or a'. We call this event E_n. According as $S_n=a$ or a', $T_{n_{k+1}}-T_n$ is the number of 0's or $(a-a')$'s in the sequence of partial sums of $X_{n+1},\cdots,X_{n_{k+1}}$. Let the event

$$|T_{n_{k+1}}-T'_{n_{k+1}}-(T_n-T'_n)| \le M_{n_{k+1}}^{3/4+\epsilon}$$

be denoted by $E_{n,n_{k+1}}$. By (10), if k is sufficiently large,

$$P(E_{n,n_{k+1}}|E_n) \ge 1 - M_{n_{k+1}}^{-\epsilon} \ge \frac{1}{2}.$$

Further it is clear that

$$P(E_{n,n_{k+1}}|E'_{n_k}\cdots E'_{n-1}E_n)^* = P(E_{n,n_{k+1}}|E_n) \ge \frac{1}{2}$$

* If E,F are two events, E' denotes the negation of E, EF denotes the conjunction of E and F.

(this follows from the Markov property of the sequence S_n.) Now $E_{n,n_{k+1}}$ and $E'_{n_k}\dots E'_{n-1}E_n$ together imply

$$|T_{n_{k+1}}-T'_{n_{k+1}}| > M_{n_{k+1}}^{3/4+\varepsilon}.$$

Hence

$$P(|T_{n_{k+1}}-T'_{n_{k+1}}| > M_{n_{k+1}}^{3/4+\varepsilon}) \geq \sum_{n=n_k}^{n_{k+1}} P(E'_{n_k}\dots E'_{n-1}E_n)P(E_{n,n_{k+1}}|E'_{n_k}\dots E'_{n-1}E_n)$$

$$\geq \frac{1}{2}\sum_{n=n_k}^{n_{k+1}} P(E'_{n_k}\dots E'_{n-1}E_n) = \frac{1}{2}P(\max_{n_k\leq n\leq n_{k+1}} |T_n-T'_n|M_n^{-3/4-\varepsilon} > 2).$$

Thus by (10)

$$\sum_k P(\max_{n_k \leq n \leq n_{k+1}} |T_n-T'_n|M_n^{-3/4-\varepsilon} > 2) \leq 2 \sum_k M_{n_{k+1}}^{-\varepsilon} < \infty.$$

It follows from the Borel-Cantelli lemma that

$$P(|T_n-T'_n| > 2M_n^{3/4+\varepsilon} \quad \text{i.o.}) = 0.$$

This is equivalent to the statement of Theorem 5.

The next theorem is a new type of limit theorem. The sequence of random variables $Y_1,Y_2,\dots$ does not obey the usual law of large numbers in the sense that constants A_n do not exist so that with probability 1,

$$\lim_{n\to\infty} \frac{Y_1+\dots+Y_n}{A_n} = 1.$$

By analogy with the situation for sums of independent random variables with finite first moments, we should expect to take A_n to be the M_n above. That this is not true is shown already in the simplest case of Bernoullian variables $X_1,X_2,\dots$ where each $X_k = \pm 1$ each with probability 1/2. In this case $m_k \sim A\, k^{-1/2}$, $M_n \sim 2A_n^{1/2}$, but the sum $Y_1+\dots+Y_n$ oscillates between $A'n^{1/2}(\log n)^{-1-\varepsilon}$ and $A''\, n^{1/2}(\log\log n)^{1/2}$ with probability 1 (see [6]). However we shall show in the next theorem that, in a certain sense, Y_k does behave like its expectation m_k, as follows.

THEOREM 6. Under (0),

$$P\left(\lim_{n \to \infty} \frac{1}{\log M_n} \sum_{k=1}^{n} \frac{Y_k}{M_k} = 1 \right) = 1.$$

Notice that in this formula if we replace Y_k by m_k, the limit relation holds without the intervention of probability. If we regard $(Y_1+\ldots+Y_n)/M_n$ as a sort of 'arithmetical average,' the quantity

$$\frac{1}{\log M_n} \sum_{k=1}^{n} \frac{Y_k}{M_k}$$

may be called a 'logarithmic average.' Evidently the existence of the mathematical average implies the existence (and equality therewith) of the logarithmic. The first instance of considering such an average in probability is due to P. Levy [9], p.270.

Proof of Theorem 6. We have

$$E\left(\sum \frac{Y_k}{M_k} \right) = \sum \frac{m_k}{M_k} = \log M_n + O(1).$$

Next

$$E\left(\left(\sum \frac{Y_k}{M_k} \right)^2 \right) = \sum \frac{m_k}{M_k^2} + 2 \sum_{j<k} \frac{m_j m_{k-j}}{M_j M_k}$$

$$= O(1) + 2 \sum_{j=1}^{n} \frac{m_j}{M_j} \sum_{k=j+1}^{n} \frac{m_{k-j}}{M_k}.$$

Hence

$$0 \le E\left(\left(\sum \frac{Y_k}{M_k} \right)^2 \right) - E^2\left(\sum \frac{Y_k}{M_k} \right)$$

$$\le O(1) + O\left(\sum \frac{m_k^2}{M_k^2} \right) + 2 \sum_{j=1}^{n} \frac{m_j}{M_j} \left(\sum_{k=j+1}^{n} \frac{m_{k-j}}{M_k} - \sum_{k=j+1}^{n} \frac{m_k}{M_k} \right)$$

$$\le O(\log M_n) + 2 \sum_{j=1}^{n} \frac{m_j}{M_j} \left\{ \sum_{k=1}^{n-j} \frac{m_k}{M_{k+j}} - \sum_{k=j+1}^{n} \frac{m_k}{M_k} \right\}$$

$$\leq O(\log M_n) + 2 \sum_{j=1}^{n} \frac{m_j}{M_j} \sum_{k=1}^{j} \frac{m_k}{M_{k+j}}$$

$$\leq O(\log M_n) + 2 \sum_{j=1}^{n} \frac{m_j}{M_j} = O(\log M_n).$$

By Chebychev inequality

$$P(\left| \sum \frac{Y_k}{M_k} - \log M_n \right| > \varepsilon \log M_n) \leq O(\frac{1}{\log M_n}).$$

Choose an increasing sequence n_k such that

$$M_{n_k} \sim e^{k^2}.$$

By the Borel-Cantelli lemma,

$$P(\lim_{k \to \infty} \frac{1}{\log M_{n_k}} \sum_{i=1}^{n_k} \frac{Y_i}{M_i} = 1) = 1.$$

Now if $n_k \leq n \leq n_{k+1}$,

$$\frac{1}{\log M_{n_{k+1}}} \sum_{k=1}^{n_k} \frac{Y_i}{M_i} \leq \frac{1}{\log n} \sum_{i=1}^{n} \frac{Y_i}{M_i}$$

$$\leq \frac{1}{\log M_{n_k}} \sum_{i=1}^{n_{k+1}} \frac{Y_i}{M_i} .$$

Since $\log M_{n_{k+1}} / \log M_{n_k} \to 1$ as $k \to \infty$ the extreme sides of these inequalities $\to 1$ with probability 1, by what has just been proved. Theorem 6 follows.

THEOREM 7. Under (0), for every $\varepsilon > 0$

$$P(M_n^{1-\varepsilon} < T_n < M_n^{1+\varepsilon} \text{ for all sufficiently large } n) = 1.$$

Proof. This is equivalent to the following two statements:

(1) $$P(T_n > M_n^{1+\varepsilon} \quad \text{i.o.}) = 0$$

(2) $$P(T_n < M_n^{1-\varepsilon} \quad \text{i.o.}) = 0.$$

The proof of (1) is similar to that of Theorem 5 and will be omitted. To prove (2), we choose $v=v(n)$ such that $M_v \sim M_n^{1-\varepsilon}$; this is possible because $m_n \longrightarrow 0$ and $M_n \uparrow \infty$. From Theorem 6 we have, with probability 1,

$$\lim_{n\to} \frac{1}{\log M_n} \sum_{k=1}^{v} \frac{Y_k}{M_k} = 1 - \varepsilon$$

$$\lim_{n\to\infty} \frac{1}{\log M_n} \sum_{k=v+1}^{n} \frac{Y_k}{M_k} = \varepsilon.$$

Upon subtraction it follows that

$$\lim_{n\to\infty} \frac{1}{\log M_n} \frac{T_n - T_v}{M_v} \geq \varepsilon$$

or

$$\lim_{n\to\infty} \frac{T_n}{M_n^{1-\varepsilon} \log M_n} \geq \varepsilon.$$

This is equivalent to (2).

Remark. Part (2) of Theorem 7 would also have followed from a general theorem of Feller (Theorem 2 in [10]), but for the condition (13) there. To verify this condition (or rather a slightly weaker one) it would be sufficient to show that

$$M_{2n} \leq 2M_n.$$

We are unable to prove or disprove this relation.

THEOREM 8. Under (0)

$$P\left(\lim_{n\to\infty} \frac{T_n}{T'_n} = 1\right) = 1.$$

Proof. This is an immediate consequence of Theorems 5 and 7. Actually we have even, for every $\varepsilon > 0$

$$P\left(\left|\frac{T_n - T'_n}{T_n}\right| > M_n^{-(1+\varepsilon)/4} \quad \text{i.o.}\right) = 0.$$

We are indebted to Dr. Miriam Lipschutz for several corrections on the MS.

Cornell University
University of Illinois

19 PROBABILITY LIMIT THEOREMS ASSUMING ONLY THE FIRST MOMENT I

BIBLIOGRAPHY

[1] B. V. Gnedenko. (a) On a local limit theorem of the theory of probability. (In Russian) Uspehi Matem. Nauk. (N.S.) 3 (1948), pp. 187-197; (b) On a local theorem for the region of normal attraction of stable laws. (In Russian) Doklady Akad. Nauk. SSSR (N.S.) 66(1949), pp. 325-326.

[2] E. R. van Kampen and A. Wintner. A limit theorem for probability distributions in lattices. Amer. J. Math. 61 (1939), pp. 965-973.

[3] C. G. Esseen. Fourier analysis of distribution functions. Acta Math. 77 (1945), pp. 1-125.

[4] K. L. Chung and W. H. J. Fuchs. On the distribution of values of sums of random variables.

[5] W. Feller. Note on the law of large numbers and 'fair' games. Ann. Math. Stat. 16 (1945), pp. 301-304.

[6] K. L. Chung and G. A. Hunt. On the zeros of $\sum_1^n \pm 1$. Ann. Math. 50(1949), pp. 385-400.

[7] W. Feller. Fluctuation theory of recurrent events. Trans. Amer. Math. Soc. 67(1949), pp. 98-119.

[8] K. L. Chung. Fluctuations of sums of independent random variables. To appear in Ann. Math.

[9] P. Lévy. Théorie de l'addition des variables aléatoires. Paris, 1937.

[10] W. Feller. A limit theorem for random variables with infinite moments. Amer. J. Math. 68(1946), pp. 257-262.

[11] W. Doeblin. Sur deux problèmes de M. Kolmogoroff concernant les chaines dénombrables, Bull. Soc. Math. de France 66(1938), pp. 210-220.

Reprinted from *Trans. Amer. Math. Soc.* 72 (1952), pp. 179–186
with permission from the American Mathematical Society.

ON THE APPLICATION OF THE BOREL-CANTELLI LEMMA

BY

K. L. CHUNG[1] AND P. ERDÖS

Consider a probability space $(\Omega, \mathcal{C}, P)$ and a sequence of events ($\mathcal{C}$-measurable sets in Ω) $\{E_k\}$, $k=1, 2, \cdots$. The upper (or outer) limiting set of the sequence $\{E_k\}$ is defined by

$$\limsup E_k = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} E_k.$$

We recall that the events E_k are said to be (mutually) independent (with respect to the probability measure P) if for any finite number of distinct subscripts $k_1, \cdots, k_s$ we have

$$P(E_{k_1} \cdots E_{k_s}) = P(E_{k_1}) \cdots P(E_{k_s}).$$

The celebrated Borel-Cantelli lemma asserts that

(A) If $\sum P(E_k) < \infty$, then P $(\limsup E_k) = 0$;

(B) If the events E_k are independent and if $\sum P(E_k) = \infty$, then $P(\limsup E_k) = 1$. In intuitive language $P(\limsup E_k)$ is the probability that the events E_k occur "infinitely often" and will be denoted by $P(E_k \text{ i.o.})$. This lemma is the basis of all theorems of the strong type in probability theory. Its application is made difficult by the assumption of independence in part (B). As Borel already noticed [1, p. 48 ff.], this assumption can be removed if we assume that[2]

$$\sum P(E_k \mid E_1' \cdots E_{k-1}') = \infty \tag{0}$$

where $P(F \mid E)$ denotes the conditional probability of F on the hypothesis of E and E' denotes the complement of E. Although Borel used the condition (0) successfully in his pioneering work on the metric theory of continued fractions, it is too stringent for many purposes. To overcome the difficulty one usually constructs a sequence of independent events out of the given sequence and applies (B) to the new one. This is the device used for instance in the proof of the law of the iterated logarithm and similar theorems. There is however another group of strong theorems to which this method does not

Received by the editors June 14, 1951.

(1) Research supported by an ONR contract.

(2) Clearly we may suppose that $P(\bigcup_{j=1}^{k} E_j) < 1$ for every k so that the conditional probability is defined. *Added in proof.* Conditions like (0) were used a great deal by Paul Lévy and other authors in generalizations to dependent variables; however, that is not what we have in mind here.

seem to apply. The following theorem furnishes an alternative method which may be of fairly general applicability. On the other hand it does not seem to apply to the law of the iterated logarithm, etc. Two examples given below will serve as illustrations, of which the second concerns the arcsin law.

THEOREM 1. *Let* $\{E_k\}$ *be a sequence of events satisfying*:

(i) $\sum P(E_k) = \infty$.

(ii) *For every pair of positive integers* h, n *with* $n \geqq h$ *there exist* $c(h)$ *and* $H(n, h) > h$ *such that for every* $k \geqq H(n, h)$ *we have*

$$P(E_k \mid E'_h \cdots E'_n) > cP(E_k). \tag{1}$$

(iii) *There exist two absolute constants* c_1 *and* c_2 *with the following property: to each* E_j *there corresponds a set of events* $E_{j_1}, \cdots, E_{j_s}$ *belonging to* $\{E_k\}$ *such that*

$$\sum_{i=1}^{s} P(E_j E_{j_i}) < c_1 P(E_j) \tag{2}$$

and if $k > j$ *but* E_k *is not among the* E_{j_i} $(1 \leqq i \leqq s)$ *then*

$$P(E_j E_k) < c_2 P(E_j) P(E_k). \tag{3}$$

Then $P(E_k \ i.o.) = 1$.

A defense of the assumptions made seems in order. The conditions (i) and (ii) together resemble Borel's condition (0) but actually they are very much weaker. The point is that the function $H(n)$ is at our disposal and can be chosen of an infinitely greater order of magnitude then n. To put it in a picturesque way, (iii) requires only that the *arbitrarily* remote past should have no overwhelming effect on the present which is certainly a state of affairs to be hoped for in probability problems. As regards the additional conditions in (iii), they involve only joint probabilities of *pairs* of events, or what is sometimes referred to as dependence to the second order; part (2) would usually deal with the dependence at close range while (3) deals with the general situation.

Before proceeding to the proof we shall state a simple lemma.

LEMMA. *Let* $\{F_k\}$, $k = 1, \cdots, N$, *be an arbitrary sequence of events in* $(\Omega, \mathcal{C}, P)$. *We have, if* $P(\bigcup_{k=1}^{n} F_k) > 0$,

$$2 \sum_{1 \leqq j < k \leqq N} P(F_j F_k) \geqq \left[P\left(\bigcup_{k=1}^{N} F_k \right) \right]^{-1} \left(\sum_{k=1}^{N} P(F_k) \right)^2 - \sum_{k=1}^{N} P(F_k). \tag{4}$$

Proof. Define random variables $X_k(\omega)$, $\omega \in \Omega$, as follows:

$$X_k(\omega) = \begin{cases} 0 & \text{if} \quad \omega \notin F_k, \\ 1 & \text{if} \quad \omega \in F_k. \end{cases}$$

The following identity is evident:

$$(5)\quad 2\sum_{1\leq j<k\leq N} P(F_jF_k) = E\{(X_1+\cdots+X_N)^2\} - E(X_1^2+\cdots+X_N^2).$$

Now by the Schwarz inequality we have

$$(6)\quad [E(X_1+\cdots+X_N)]^2 \leqq P(X_1+\cdots+X_N>0)E\{(X_1+\cdots+X_N)^2\}.$$

Since $E(X_k)=E(X_k^2)=P(F_k)$, $P(X_1+\cdots+X_N>0)=P(\bigcup_{k=1}^N F_k)$ by definition, (4) follows from (5) and (6).

Proof of Theorem 1. Let

$$B_h = \bigcup_{k=h}^{\infty} E_k.$$

Since $(E_k \text{ i.o.}) = \lim_{h\to\infty} P(B_h)$, it is sufficient to prove that $P(B_h)=1$ for every h. Suppose that this is not true for a certain h; let $P(B_h)=1-\delta$, $\delta>0$. Thus

$$(7)\qquad P\left(\bigcap_{k=h}^{\infty} E_k'\right) = \delta > 0.$$

Given any ϵ, $0<\epsilon<1-\delta$, we can find n such that $P(\bigcup_{k=h}^n E_k)>1-\delta-\epsilon$ so that if we write $D_{h,n}=\bigcup_{k=h}^{\infty} E_k - \bigcup_{k=h}^n E_k$, we have

$$(8)\qquad P(D_{h,n}) < \epsilon.$$

We have by (1) and (7), if $k>H(n)$,

$$(9)\qquad P(E_kE_h'\cdots E_n') > c\delta P(E_k).$$

Hence by (i), $\sum_{k=H(n)}^{\infty} P(E_kE_h'\cdots E_n')=\infty$. Therefore there exists an integer $H'(n)>H(n)$ such that $(H=H(n),\ H'=H'(n))$

$$(10)\qquad 1 < \sum_{k=H}^{H'} P(E_kE_h'\cdots E_n') \leqq 2.$$

From (9) and (10) we obtain

$$(11)\qquad \sum_{k=H}^{H'} P(E_k) < \frac{2}{c\delta}\cdot$$

From (2), (3), and (11) we have

$$(12)\qquad \begin{aligned}\sum_{H\leq j<k\leq H'} P(E_jE_k) &\leqq c_1\sum_{j=H}^{H'} P(E_k) + c_2 \sum_{H\leq j<k\leq H'} P(E_j)P(E_k)\\ &< \frac{2c_1}{c\delta}+\frac{c_2}{2}\left(\frac{2}{c\delta}\right)^2 = c_3(\delta),\end{aligned}$$

where $c_3(\delta)$ is a constant defined by the last equality.

Now let $F_k = E_k E_k' \cdots E_n'$, $H \leqq k \leqq H'$. It is obvious that $\bigcup_{k=H}^{H'} F_k$ is a subset of $D_{h,n}$, hence by (8),

$$P\left(\bigcup_{k=H}^{H'} F_k\right) < \epsilon. \tag{13}$$

From (10) and (11) we have

$$1 < \sum_{k=H}^{H'} P(F_k) \leqq \sum_{k=H}^{H'} P(E_k) < \frac{2}{c\delta} \cdot \tag{14}$$

Applying the lemma to $\{F_k\}$, $H \leqq k \leqq H'$, we obtain using (13) and (14)

$$2 \sum_{H \leqq j < k \leqq H'} P(E_j E_k) \geqq 2 \sum_{H \leqq j < k \leqq H'} P(F_j F_k) \geqq \frac{1}{\epsilon} - \frac{2}{c\delta} \cdot \tag{15}$$

But (12) and (15) are incompatible for sufficiently small ϵ. This contradiction proves that $\delta = 0$. Hence $P(B_h) = 1$. q.e.d.

In the two applications given below we shall treat only the simplest Bernoullian case, since we are more interested in the principle involved than the technical difficulties. It is not hard to generalize Theorems 2 and 3 to fairly general lattice cases or even continuous cases. It will be seen from their proofs that only certain asymptotic formulas and a kind of boundedness of S_n, with probability one or even in probability, are required. These are available in more general cases, thanks to various modern limit theorems.

THEOREM 2. *Let $\{X_k\}$, $k = 1, 2, \cdots$, be independent random variables and each X_k assume the values $+1$ and -1 with probabilities $1/2$ and $1/2$. Let $S_n = \sum_{k=1}^n X_k$. Let $\{n_i\}$, $i = 1, 2, \cdots$, be an increasing sequence of even integers such that there exists an absolute constant A with the property that*

$$n_{i+1} - n_i > A n_i^{1/2}. \tag{16}$$

Then

$$P(S_{n_i} = 0 \ i.o.) = \begin{cases} 0 \\ 1 \end{cases}$$

according as

$$\sum_i n_i^{-1/2} \begin{Bmatrix} < \\ = \end{Bmatrix} \infty.$$

REMARK. The theorem in the divergence case is not true without some such condition as (16). Example: Take $\{n_i\}$ to be the sequence of even integers in the intervals $[k^8, k^8 + k^5]$, $k = 1, 2, \cdots$. For an alternative condi-

tion and proof see [3, p. 1009].

Proof. The convergence case follows directly from part (A) of the Borel-Cantelli lemma without the condition (16).

Next, let E_i denote the event $S_{n_i}=0$. We know that

$$P(E_i) = P(S_{n_i} = 0) \sim (2/\pi n_i)^{1/2}.$$

Hence condition (i) in Theorem 1 is satisfied.

To verify the condition (ii) in Theorem 1 we notice that $|S_{n_i}| \leqq n_i$, hence

$$\begin{aligned} &P(S_{n_k} = 0 \mid S_{n_h} \neq 0, \cdots, S_{n_i} \neq 0) \\ &\qquad \geqq \operatorname*{Min}_{|x| \leqq n_i} P(S_{n_k} = 0 \mid S_{n_i} = x) = \operatorname*{Min}_{|x| \leqq n_i} P(S_{n_k} - S_{n_i} = -x) \\ &\qquad = \operatorname*{Min}_{|x| \leqq n_i} P(S_{n_k - n_i} = -x). \end{aligned}$$

Now we have, if $x^2 = o(n)$,

$$P(S_n = x) = C\binom{n}{(n-x)/2}\frac{1}{2^n} \sim \left(\frac{2}{\pi n}\right)^{1/2} e^{-x^2/2n} \sim \left(\frac{2}{\pi n}\right)^{1/2}.$$

We choose $H(i)$ sufficiently large so that if $k > H(i)$, then $n_i^2 = o(n_k)$. Then we have for all $|x| \leqq n_i$,

$$P(S_{n_k - n_i} = -x) \sim \left(\frac{2}{\pi n_k}\right)^{1/2}.$$

Therefore we have for all h, $i \geqq h$ and $k \geqq H(i)$ and any fixed $c < 1$, if $H(1)$ is sufficiently large,

$$P(S_{n_k} = 0 \mid S_{n_h} \neq 0, \cdots, S_{n_i} \neq 0) > cP(S_{n_k} = 0).$$

Thus condition (ii) in Theorem 1 is satisfied.

To verify the condition (iii) in Theorem 1 we have

$$\begin{aligned} P(E_j E_k) &= P(S_{n_j} = 0) P(S_{n_k - n_j} = 0) \\ &\sim P(S_{n_j} = 0)\left(\frac{2}{\pi(n_k - n_j)}\right)^{1/2} \\ &\sim P(S_{n_j} = 0) P(S_{n_k} = 0)\left(\frac{n_k}{n_k - n_j}\right)^{1/2} \qquad (17) \\ &= P(E_j) P(E_k)\left(\frac{n_k}{n_k - n_j}\right)^{1/2}. \end{aligned}$$

If $n_k > 2n_j$, then

$$\left(\frac{n_k}{n_k - n_j}\right)^{1/2} < 2^{1/2}. \qquad (18)$$

We call the events E_k with $n_j < n_k \leqq 2n_j$ the events E_{j_i} $(1 \leqq i \leqq s)$ associated with each E_j. We have as before

$$(19) \qquad \sum{}' P(E_j E_k) \sim P(E_j) \sum{}' (n_k - n_j)^{-1/2}$$

where the summation extends to those k for which $n_j < n_k \leqq 2n_j$. From (16) we deduce that if $k > j$ (A_1 denoting an absolute constant),

$$n_k - n_j > A_1(k^2 - j^2).$$

Let N denote the number of k's satisfying $n_j < n_k \leqq 2n_j$. From the last inequality we deduce that $n_j + A_1(N^2 + 2jN) \leqq n_{j+N} \leqq 2n_j$. Hence we have

$$(20) \qquad N \leqq \left(\frac{n_j}{A_1}\right)^{1/2}.$$

Now using the Schwarz inequality, (16), and (20) we obtain

$$(21) \quad \left(\sum{}' (n_k - n_j)^{-1/2}\right)^2 \leqq N \sum{}' (n_k - n_j)^{-1} \leqq \left(\frac{n_j}{A_1}\right)^{1/2} \frac{1}{An_j^{1/2}} = \left(\frac{1}{A_1 A^2}\right)^{1/2}.$$

(19) and (21) give (2) while (17) and (18) give (3). q.e.d.

THEOREM 3. *Let $\{X_n\}$ be as in Theorem 2 and let N_n denote the number of positive terms among $S_1, \cdots, S_n$. Let $\phi(n)$ be an increasing function of n. Then*

$$(22) \qquad P\left(N_n \leqq \frac{n}{\phi(n)} \text{ i.o.}\right) = \begin{cases} 0 \\ 1 \end{cases}$$

according as

$$(23) \qquad \sum \frac{1}{n(\phi(n))^{1/2}} \begin{Bmatrix} < \\ = \end{Bmatrix} \infty.$$

REMARK. This is the strong theorem corresponding to the now celebrated arcsin law. On grounds of symmetry we may replace the left side in (22) by $P(N_n \geqq n(1 - 1/\phi(n)) \text{i.o.})$.

Proof. Standard arguments(3) show that we may suppose that $\phi(n) \leqq n^\epsilon$ for some $0 < \epsilon < 1/2$. The convergence case follows easily from the arcsin law for Bernoullian variables (see [2, p. 252]; the convention made there regarding the "positiveness" of S_n makes no difference in the asymptotic formula below), which asserts that

$$P\left(N_n \leqq \frac{n}{\phi(n)}\right) \sim \frac{2}{\pi(\phi(n))^{1/2}}.$$

To prove the theorem in the divergence case we note first that the di-

(3) Cf e.g. [3. p. 1010].

vergence of the series in (23) implies that of

$$\sum_n \frac{1}{n(\phi(n^r))^{1/2}}$$

for any $r>0$ (proof by the integral test). Let $\phi(n^2)=\psi(n)$. Define E_k to be the event

$$S_{2k}=0, \qquad S_i<0 \qquad \text{for } 2k<i\leqq 2k\psi(k).$$

Obviously E_k implies that $N_{2k\psi(k)}\leqq 2k$. Writing $2k\psi(k)=n$, we have, since $\psi(k)\leqq k^{2\epsilon}$, $k\geqq n'$ where $n'=(2^{-1}n)^{1/(1+2\epsilon)}$. Hence $N_n\leqq n/\psi(n')$. For all sufficiently large n, $\psi(n')\geqq\phi(n)$. Hence in order to prove the second part in (22) it is sufficient to prove that $P(E_k \text{ i.o.})=1$.

It is known that (see e.g. [2, p. 252]) $P(S_i<0 \text{ for } 0<i\leqq n)\sim bn^{-1/2}$ for some absolute constant $b>0$. Hence we have

$$\begin{aligned}(24)\qquad P(E_k) &= P(S_{2k}=0)P(S_i<0 \text{ for } 0<i\leqq 2k\psi(k)-2k)\\ &\sim bk^{-1/2}(k\psi(k))^{-1/2}=bk^{-1}(\psi(k))^{-1/2}.\end{aligned}$$

Hence condition (i) in Theorem 1 is satisfied.

To verify condition (ii) in Theorem 1 we note that (without loss of generality we may suppose $n\psi(n)$ to be an integer for all n), if $k>H(n)$,

$$\begin{aligned}P(E_k\mid S_1&=x_1,\cdots,S_{2n\psi(n)}=y)\\ &=\sum_x P(E_k\mid S_{H(n)}=x)P(S_{H(n)}=x\mid S_{2n\psi(n)}=y)\\ &=\sum_x P(E_k\mid S_{H(n)}=x)P(S_{H(n)-2n\psi(n)}=x-y)\end{aligned}$$

where $x_1,\cdots,y,x$ are integers. Now $|y|\leqq 2n\psi(n)$, hence if we choose $H(n)$ sufficiently large, $P(S_{H(n)-2n\psi(n)}=x-y)\sim P(S_{H(n)}=x)$ as $n\to\infty$, at least if x is within a certain range, say $|x|\leqq H(n)^{1/2+\eta}$, $\eta>0$. (This is because of the limitations of the Gaussian approximation.) But the other range of x is negligible in the sense that

$$\sum_{|x|>H(n)^{1/2+\eta}} P(S_{H(n)-2n\psi(n)}=x-y)$$
$$=O\Bigg(\sum_{|x|\leqq H(n)^{1/2+\eta}} P(S_{H(n)-2n\psi(n)}=x-y)\Bigg).$$

Hence we have

$$\underset{|y|\leqq 2n\psi(n)}{\mathrm{Min}}\ P(E_k\mid S_1=x_1,\cdots,S_{2n\psi(n)}=y)$$
$$\sim\sum_x P(E_k\mid S_{H(n)}=x)P(S_{H(n)}=x)=P(E_k).$$

This implies condition (ii) in Theorem 1.

To verify condition (iii) in Theorem 1, let $j<k$. If $k\leqq j\psi(j)$, then $P(E_jE_k)=0$. If $k>j\psi(j)$ we have

$$(25)\quad \begin{aligned} P(E_k \mid E_j) &= P(S_{2k}=0 \mid S_{2j}=0, S_i<0 \text{ for } 2j<i\leqq 2j\psi(j)). \\ & P(S_i<0 \text{ for } 2k<i\leqq 2k\psi(k) \mid S_{2k}=0) = P_1\cdot P_2. \end{aligned}$$

Now for every x we have

$$P(S_{2k}=0 \mid S_{2j\psi(j)}=x) = P(S_{2k-2j\psi(j)}=-x) \leqq b(k-j\psi(j))^{-1/2}.$$

P_1 being a probability mean of such probabilities we have

$$P_1 \leqq b(k-j\psi(j))^{-1/2}.$$

As for P_2 we have as in (24),

$$P_2 \sim b(2k\psi(k)-2k)^{-1/2} \sim b(2k\psi(k))^{-1/2}.$$

Therefore we obtain from (25),

$$(26)\quad P(E_jE_k) \leqq b_1P(E_j)(k-j\psi(j))^{-1/2}(k\psi(k))^{-1/2},$$

where b_1 (as b_2, b_3 later) is an absolute constant. Now for every E_j we define E_{j_i}, $1\leqq i\leqq s$, to be those E_k with $j\psi(j)<k\leqq 2j\psi(j)$. We have then by (26)

$$\begin{aligned} \sum_{i=1}^{s} P(E_jE_{j_i}) &\leqq b_1P(E_j)(k\psi(k))^{-1/2}\sum_{i=1}^{j\psi(j)} i^{-1/2} \\ &\leqq b_2P(E_j). \end{aligned}$$

On the other hand if $k>2j\psi(j)$, then $k-j\psi(j)>k/2$, hence by (26) and (24)

$$P(E_jE_k) \leqq b_3P(E_j)P(E_k).$$

Therefore condition (iii) in Theorem 1 is satisfied. q.e.d.

References

1. E. Borel, *Traité du calcul des probabilités et de ses applications*, vol. 2, no. 1, *Applications à l'arithmétique et à la théorie des fonctions*, Paris, Gauthier-Villars, 1926.

2. W. Feller, *An introduction to probability theory and its applications*, vol. 1, New York, Wiley, 1950.

3. K. L. Chung and P. Erdös, *On the lower limit of sums of independent random variables*, Ann. of Math. vol. 48 (1947) p. 1003.

Columbia University,
New York, N. Y.
University of Aberdeen,
Aberdeen, Scotland.

Journal of Research of the National Bureau of Standards Vol. 50, No. 4, April 1953 Research Paper 2411

Contributions to the Theory of Markov Chains

Kai Lai Chung [1]

The fundamentals of the theory of denumerable Markov chains with stationary transition probabilities were laid down by Kolmogorov, and further work was done by Doblin. The theory of recurrent events of Feller is closely related, if not coextensive. Some new results obtained by T. E. Harris turn out to tie up very nicely with some amplifications of Doblin's work. Harris was led to consider the probabilities of hitting one state before another, starting from a third one. This idea of considering three states, one initial, one "taboo", and one final, is more fully developed in the present work. The notion of first passage time to the "union" or "intersection" of two states is also introduced here. The interplay between these notions is illustrated.

The fundamentals of the theory of denumerable Markov chains [2] with stationary transition probabilities (DMCS) were laid down by Kolmogorov [1] [3] and further work was done by Doblin [2]. The theory of recurrent events of Feller [3] is closely related, if not coextensive. Recently some interesting new results were obtained by T. E. Harris [4] and communicated to the author. They turn out to tie up very nicely with some amplifications of Doblin's work the author was engaged in. Although Harris' main purpose lies elsewhere, he was led to consider the probabilities of hitting one state before another, starting from a third one. This idea of considering three (instead of the customary two) states, one initial, one "taboo," and one final, will be more fully developed in the present work. The notion of first passage time to the "union" or "intersection" of two states will also be introduced here. The interplay between these notions will be illustrated.

Recorded results in this paper will be labeled as formulas and theorems, respectively. Relevant remarks as to their origin or significance will be found in the body of the paper. The author is indebted to Dr. Harris for communicating some of his results before publication.

1. The sequence of random variables $\{X_n\}$, $n=0,1,2,\ldots$ forms a DMCS. The states will be denoted simply by the positive integers. The (one-step) transition probability from the state i to the state j will be denoted by $P^{(1)}_{ij}$. Because of stationarity we have

$$P^{(1)}_{ij}=P(X_{m+1}=j|X_m=i)$$

for all integers $m\geq 0$ for which the conditional probability is defined. With this understanding, we shall permit ourselves to write $m=0$ in the definitions to follow, as if the conditional probabilities were always defined.

NOTATIONS:

n, N, v, r, s, denote positive integers and will be used as time parameters or general numerals;

i, j, k, l, h, denote positive integers and will be used as state labels:

$$P^{(n)}_{ij}=P(X_n=j|X_0=i); \qquad P^{(0)}_{ij}=\begin{cases}0, & i\neq j\\ 1, & i=j\end{cases}$$

$${}_kP^{(n)}_{ij}=P(X_n=j,\quad X_v\neq k,\quad 1\leq v<n|X_0=i)$$

$$F^{(n)}_{ij}=P(X_n=j,\quad X_v\neq j,\quad 1\leq v<n|X_0=i)$$

$${}_kF^{(n)}_{ij}=P(X_n=j,\quad X_v\neq j,\neq k,\quad 1\leq v<n|X_0=i)$$

$$Q^*=\sum_{n=1}^{\infty}Q^{(n)}$$

where Q may stand for any of the symbols ${}_kP_{ij}$, F_{ij}, or ${}_kF_{ij}$.

We offer the following clue to the above notations. The letter P designates "passage"; the letter F, "first passage"; the first right-hand subscript designates the initial state; the second, the final state; the left-hand subscript designates the "taboo state," namely, one to be eschewed during the passage (exclusive of both terminals); the star on a letter with subscripts designates the sum of the corresponding infinite series (finite or $+\infty$) summed from $n=1$ *ad inf.* We admit that this is not the most logical system of notations we could have invented. For instance, we have the superfluity $F^{(n)}_{ij}={}_jP^{(n)}_{ij}$, and if we had allowed more than one left-hand subscript,[4] we could have used only one letter P and written ${}_kF^{(n)}_{ij}={}_{j,k}P^{(n)}_{ij}$. However, we consider our notations to be preferred to the arbitrary use of all sorts of letters from the Latin and Greek alphabets. Also, after painful deliberations we decided not to define ${}_kP^{(0)}_{ij}$, $F^{(0)}_{ij}$, or ${}_kF^{(0)}_{ij}$, while reserving the right to do so later in some cases.

FORMULA I: *If $i\neq j$, then*

$$F^*_{ij}={}_iF^*_{ij}(1+{}_jP^*_{ii}), \tag{I}$$

where on the right side $0\cdot\infty$ is to be taken as 0.[5]

[1] National Bureau of Standards and Institute of Statistics, University of North Carolina, Chapel Hill, N. C.

[2] "Denumerable" means "with a denumerable number of states;" "chain" refers to a process with an integral time parameter.

[3] Figures in brackets indicate the literature references at the end of this paper.

[4] This naturally suggests the consideration of more than one taboo state.

[5] This follows also from the easily interpreted relations

$$1+{}_jP^*_{ii}=\sum_{n=0}^{\infty}({}_jF^*_{ii})^n=\frac{1}{1-{}_jF^*_{ii}}=\frac{F^*_{ij}}{{}_iF^*_{ij}}.$$

The convention that $0\cdot\infty$ is to be taken as 0 will be understood in similar circumstances.

Proof: We start from the formula

$$F_{ij}^{(n)}=\sum_{v=0}^{n-1}{}_jP_{ii}^{(v)}\,{}_iF_{ij}^{(n-v)}, \tag{1}$$

where we agree that ${}_jP_{ii}^{(0)}=1$. Equation (1) is proved as follows. Either the state i is not entered at all during the passage from i to j, which contingency contributes the term corresponding to $v=0$ on the right side of (1); or there is a last entry of i, occurring at the vth step, $1\le v\le n-1$, which contingency contributes the general term.

Summing (1) over n, we obtain

$$1\ge F_{ij}^*=\sum_{n=1}^{\infty}F_{ij}^{(n)}=\sum_{n=1}^{\infty}\sum_{v=0}^{n-1}{}_jP_{ii}^{(v)}\,{}_iF_{ij}^{(n-v)}$$

$$=\sum_{v=0}^{\infty}{}_jP_{ii}^{(v)}\sum_{n=v+1}^{\infty}{}_iF_{ij}^{(n-v)}={}_iF_{ij}^*(1+{}_jP_{ii}^*).$$

Since the terms of the double series are nonnegative, the inversion is justified and (I) is proved. Moreover, this proves that if ${}_iF_{ij}^*>0$, then ${}_jP_{ii}^*<\infty$. It follows from (1) that ${}_iF_{ij}^*=0$ if, and only if, $F_{ij}^*=0$, namely, $P_{ij}^{(n)}=0$ for all n.

Formula II: *If $j\neq k$, then*

$${}_kP_{ij}^*={}_kF_{ij}^*(1+{}_kP_{jj}^*). \tag{II}$$

(This formula is easily interpreted in terms of mathematical expectations.)

Proof: We start from the formula

$${}_kP_{ij}^{(n)}=\sum_{v=1}^{n}{}_kF_{ij}^{(v)}\,{}_kP_{jj}^{(n-v)}, \tag{2}$$

where as before ${}_kP_{jj}^{(0)}=1$. If we ignore the left-hand subscripts, (2) reduces to a familiar formula. The proof of the latter extends immediately to (2).

Summing (2) over n we obtain

$${}_kP_{ij}^*=\sum_{n=1}^{\infty}{}_kP_{ij}^{(n)}=\sum_{n=1}^{\infty}\sum_{v=1}^{n}{}_kF_{ij}^{(v)}\,{}_kP_{jj}^{(n-v)}$$

$$=\sum_{v=1}^{\infty}{}_kF_{ij}^{(v)}\sum_{n=v}^{\infty}{}_kP_{jj}^{(n-v)}={}_kF_{ij}^*(1+{}_kP_{jj}^*).$$

We note the following corollaries to (I) and (II), to be used later.

Formula IIa: *If $i\neq j$, then*

$${}_iP_{ij}^*={}_iF_{ij}^*(1+{}_iP_{jj}^*). \tag{IIa}$$

Formula IIb: *If $j\neq k$, then*

$${}_jF_{ik}^*\,{}_kP_{ij}^*=F_{ik}^*\,{}_kF_{ij}^*. \tag{IIb}$$

Formula IIc: *If $i\neq j$, then*

$$F_{ij}^*(1+{}_iP_{jj}^*)={}_iP_{ij}^*(1+{}_jP_{ii}^*). \tag{IIc}$$

Formula III: *If $i\neq j$, then*

$$\lim_{N\to\infty}\frac{\sum_{n=0}^{N}P_{ij}^{(n)}}{\sum_{n=0}^{N}P_{ii}^{(n)}}={}_iP_{ij}^*<\infty. \tag{III}$$

Proof: We start from the formula

$$P_{ij}^{(n)}=\sum_{v=0}^{n}P_{ii}^{(v)}\,{}_iP_{ij}^{(n-v)}, \tag{3}$$

where we agree that ${}_iP_{ij}^{(0)}=0$. The proof of (3) is entirely similar to that of (1).

Summing (3) from $n=0$ to $n=N$, we obtain

$$\sum_{n=0}^{N}P_{ij}^{(n)}=\sum_{n=0}^{N}\sum_{v=0}^{n}P_{ii}^{(v)}\,{}_iP_{ij}^{(n-v)}=\sum_{v=0}^{N}P_{ii}^{(v)}\sum_{n=0}^{N-v}{}_iP_{ij}^{(n)}. \tag{4}$$

We need an elementary lemma which is frequently useful in such connections.

Lemma. Let $0\le a_v\le 1$, $b_v\ge 0$; $\sum_{v=0}^{\infty}a_v>0$, $\lim_{v\to\infty}b_v=B\le+\infty$. Then

$$\lim_{N\to\infty}\frac{\sum_{v=0}^{N}a_vb_{N-v}}{\sum_{v=0}^{N}a_v}=B.$$

Applying the lemma to (4) we obtain (III). That ${}_iP_{ij}^*<\infty$ is clear from (IIa), and the remarks at the end of the proof of (I).

Theorem 1. *The limit*

$$\lim_{N\to\infty}\frac{\sum_{n=0}^{N}P_{ii}^{(n)}}{\sum_{n=0}^{N}P_{jj}^{(n)}} \tag{5}$$

exists, and is equal to any of the following three expressions:

$$\frac{1+{}_jP_{ii}^*}{1+{}_iP_{jj}^*},\quad \frac{F_{ji}^*}{{}_iP_{ij}^*},\quad \frac{F_{ij}^*\,{}_jF_{ji}^*}{F_{ji}^*\,{}_iF_{ij}^*}; \tag{IVa, b, c}$$

the first always, the second if $i\neq j$, the third if $F_{ij}^*F_{ji}^*>0$.

Proof. Doblin [2] has shown, trivially, that

$$\lim_{N\to\infty}\frac{\sum_{n=0}^{N}P_{ij}^{(n)}}{\sum_{n=0}^{N}P_{jj}^{(n)}}=F_{ij}^*. \tag{6}$$

Comparing (III) and (6), we obtain (IVb) if $i\neq j$. (IVa) now follows from (IIc) and obviously holds for $i=j$. If $F_{ij}^*F_{ji}^*>0$, then the denominator of (IVc) is not zero, and this is then equal to (IVb) by (IIb), with $k=i$.

That the limit (5) exists, and is finite and not zero, was proved by Doblin [2]; that it is equal to (IVa) was previously proved by the author [5]. The present approach seems to be the simplest.

COROLLARY. *If i,j,k,l, are distinct states of one class,*[6] *then*

$$\lim_{N\to\infty}\frac{\sum_{n=1}^{N}P_{ij}^{(n)}}{\sum_{n=1}^{N}P_{kl}^{(n)}}=\frac{F_{ij}^{*}(1+{}_{l}P_{jj}^{*})}{F_{kl}^{*}(1+{}_{j}P_{ll}^{*})}.$$

Naturally there are other expressions for it, and we omit the tedious considerations when some of the states are identical.

2. We now consider two states i and j belonging to the same recurrent class, namely:

$$F_{ii}^{*}=F_{jj}^{*}=F_{ij}^{*}=1.$$

A fundamental idea in the theory of DMCS, already found in Kolmogorov's work, is that whatever transpires between successive entries at a recurrent state forms a sequence of independent events. Using this idea, Harris [4] and Lévy [10], independently of each other, discovered theorem 2. Our proof is somewhat different from theirs.

Let $i\neq j$ and define

$Y_n=$ the number of v, $1\leq v\leq n$, such that $X_v=i$;

$Z_n=$ the number of v, $1\leq v\leq n$, such that $X_v=j$.

In words, Y_n (or Z_n) is the number of entries at the state i (or j) in the first n steps. Using the average ergodic theorem (see (11) below) it is easy to show that if j is a positive state, and $P(X_0\epsilon C)=1$, where C is the class containing j, then we have

$$P\left(\lim_{n\to\infty}\frac{Z_n}{\sum_{v=0}^{n}P_{jj}^{(v)}}-1\right)=1.$$

The following theorem covers both positive and null classes.

THEOREM 2 (Harris-Lévy). *If i and j are two states in a recurrent class C and $P(X_0\epsilon C)=1$, then*

$$P\left[\lim_{n\to\infty}\frac{Z_n}{Y_n}=\lim_{n\to\infty}\frac{\sum_{v=0}^{n}P_{jj}^{(v)}}{\sum_{v=0}^{n}P_{ii}^{(v)}}\right]=1. \tag{7}$$

PROOF. Since i is recurrent, we have $P\left(\lim_{n\to\infty}Y_n=+\infty\right)=1$. Let $v_1<v_2<\ldots$ be the successive indices v, such that $X_v=i$. Let $W_s=$ the number of v, $v_s<v<v_{s+1}$ such that $X_v=$j. Then as remarked above, the W_s's are independent, identically distributed random variables. Evidently we have

$$E(W_s)=\sum_{n=1}^{\infty}{}_{i}P_{ij}^{(n)}={}_{i}P_{ij}^{*}<\infty.$$

Now by definition we have $v_{Y_n}\leq n<v_{Y_n+1}$ and

$$\sum_{s=1}^{Y_n-1}W_s\leq Z_N\leq v_1+\sum_{s=1}^{Y_n}W_s.$$

Consequently,

$$\frac{\sum_{s=1}^{Y_n-1}W_s}{Y_n}\leq\frac{Z_n}{Y_n}\leq\frac{v_1}{Y_n}+\frac{\sum_{s=1}^{Y_n}W_s}{Y_n}. \tag{8}$$

Applying Khintchine-Kolmogorov's strong law of large numbers (see, e. g., [9] p. 208) to the sequence $\{W_s\}$ we obtain:

$$P\left(\lim_{Y_n\to\infty}\frac{\sum_{s=1}^{Y_n}W_s}{Y_n}={}_{i}P_{ij}^{*}\right)=1. \tag{9}$$

Moreover, $P(v_1<+\infty)=1$. It follows from (8) and (9) that

$$P\left(\lim_{n\to\infty}\frac{Z_n}{Y_n}={}_{i}P_{ij}^{*}\right)=1. \tag{10}$$

Now $F_{ij}^{*}=1$. Hence theorem 2 follows from (10) and theorem 1, using (IVb) there.

This theorem includes as special case a previous result by Erdös and the author [7]. Consider independent, identically distributed random variables $\{U_n\}$ which assume only integer values with mean zero. They form a DMCS with all integers as the states. Since the mean is zero, all possible states are recurrent by a theorem of Fuchs and the author [8].[7] Without loss of generality, we may suppose that every integer is a possible, therefore recurrent, state.

Writing $S_n=\sum_{v=1}^{n}U_v$, we see that

$$P_{ij}^{(n)}=P(S_n=j-i).$$

Hence, $P_{ii}^{(n)}=P_{jj}^{(n)}=P_{00}^{(n)}$, and (7) becomes

$$P\left(\lim_{n\to\infty}\frac{Z_n}{Y_n}=1\right)=1,$$

which is theorem 8 in [7]. Needless to say, as far as this statement is concerned, Harris' approach is incomparably better. However, we note that there we actually proved a sharper result, i. e.

$$P\left(\frac{|Z_n-Y_n|}{Y_n}>M_n^{-\frac{1+\epsilon}{4}}\text{ i.o.}\right)=0$$

[6] Slightly generalizing Kolmogorov, we define a class to be a set of states such that for any two states i and j belonging to it, we have $F_{ij}^{*}F_{ji}^{*}>0$. See [6].

[7] This important step cannot be circumvented by the present, more general, approach.

for every $\epsilon>0$, where $M_n=\sum_{v=1}^{n} P(S_v=i)$. See also theorem 7 in [7]. It would be of interest to investigate corresponding strong relations for the general Markov-chain case, using perhaps a more precise form of the strong law of large numbers.

3. We now consider a positive recurrent class C. According to Kolmogorov, in C all mean recurrence and first passage times are finite, namely, for all $i,j\in C$ we have

$$m_{ij}=\sum_{n=1}^{\infty} nF_{ij}^{(n)}<\infty.$$

We introduce the notions of first passage to $j\cup k$ and to $j\cap k$, as follows. Let $j\neq k$.

$T(i,j\cup k)$ = the smallest integer $n\geq 1$ such that $X_n=j$ or $X_n=k$, whichever happens first, given that $X_0=i$;

$T(i,j\cap k)$ = the smallest integer $n\geq 1$ such that there exist two integers n_1 and n_2 such that $n_1\leq n_2$, $1\leq n_1\leq n$, $1\leq n_2\leq n$, and $X_{n_1}=j$, $X_{n_2}=k$, given that $X_0=i$;

$m(i,j\cup k)=E\{T(i,j\cup k)\}$;

$m(i,j\cap k)=E\{T(i,j\cap k)\}$.

Let w denote the sample point. Put

$$e_n^j=\{w:X_0(w)=i,\ X_v(w)\neq j,\neq k \quad \text{for } 1\leq v\leq n; X_n(w)=j\};$$

$$e^j=\bigcup_{n=1}^{\infty} e_n^j.$$

Thus e^j is the event that $X_0=i$ and the state j is reached before the state k. Since i, j, and k belong to one recurrent class, we have

$$e^j\cup e^k=\{w:X_0(w)=i\}.$$

Let $P(X_0=i)=c>0$. We have the following relations, immediate consequences of the definitions.

$$cm(i,j\cup k)=\sum_{n=1}^{\infty}\left\{\int_{e_n^j} nP(dw)+\int_{e_n^k} nP(dw)\right\}$$

$$cm_{ij}=\sum_{n=1}^{\infty}\left\{\int_{e_n^j} nP(dw)+\int_{e_n^k}(n+m_{kj})dP(w)\right\}$$

$$=cm(i,j\cup k)+P(e^k)m_{kj}$$

$$cm(i,\cap jk)=\sum_{n=1}^{\infty}\left\{\int_{e_n^j}(n+m_{jk})P(dw)+\int_{e_n^k}(n+m_{kj})P(dw)\right\}$$

$$=cm(i,j\cup k)+P(e^j)m_{jk}+P(e^k)m_{kj}.$$

Now by definition we have

$$\frac{P(e^j)}{c}={}_kF_{ij}^*,\qquad \frac{P(e^k)}{c}={}_jF_{ik}^*.$$

Hence we obtain from the above:

FORMULA VI. *If $j\neq k$, then*

$$m(i,j\cup k)=m_{ij}-{}_jF_{ik}^*m_{kj}=m_{ik}-{}_kF_{ij}^*m_{jk}. \tag{VI}$$

FORMULA VII. *If $j\neq k$, then*

$$m(i,j\cap k)=m_{ij}+{}_kF_{ij}^*m_{jk}=m_{ik}+{}_jF_{ik}^*m_{kj}. \tag{VII}$$

Since

$${}_kF_{ij}^*+{}_jF_{ik}^*=1$$

we deduce from (VII):

FORMULA VIII. *If $j\neq k$, then*

$$m_{ik}+m_{kj}-m_{ij}={}_kF_{ij}^*(m_{jk}+m_{kj}). \tag{VIII}$$

We note the following special case ($i=k$) of (VIII):

$$m_{kk}={}_kF_{kj}^*(m_{jk}+m_{kj}). \tag{VIIIa}$$

This last formula is due to Harris [4], who also derived from it the following relation:

$$\frac{m_{jj}}{m_{kk}}=\frac{{}_jF_{jk}^*}{{}_kF_{kj}^*}. \tag{VIIIb}$$

Now in a positive class the ergodic theorem of Kolmogorov holds:[8]

$$\lim_{n\to\infty}\frac{1}{n}\sum_{v=1}^{n}P_{jj}^{(v)}=\frac{1}{m_{jj}}. \tag{11}$$

Thus (VIIIb) turns out to be a special case of theorem 1, using (IVc) and noting that $F_{jk}^*=F_{kj}^*=1$.

Dividing (VIII) by the product of (VIIIa) and (VIIIb) we obtain

$$\frac{m_{ik}+m_{kj}-m_{ij}}{m_{jj}}=\frac{{}_kF_{ij}^*}{{}_jF_{jk}^*}. \tag{12}$$

By formula (IIb), the right side is ${}_kP_{ij}^*$, since $F_{jk}^*=1$ in the present case. Thus we obtain

FORMULA (IX). *If $j\neq k$ then*

$${}_kP_{ij}^*=\frac{m_{ik}+m_{kj}-m_{ij}}{m_{jj}}. \tag{IX}$$

As an application consider, as in the Central Limit Theorem for Markov chains, random variables $\{Y_n\}$ attached to the Markov chain $\{X_n\}$ in the following way: $Y_n=x_i$ if $X_n=i$ where the x_i's are arbitrary real numbers.

THEOREM 3. *Let i be a positive state. Given $X_0=i$, let v_0 denote the smallest $n\geq 1$ such that $X_n=i$. Then if the series on the right-hand side converges absolutely,*

[8] This theorem actually establishes the limit of $P_{jj}^{(n)}$ as $n\to\infty$. The average form (11) is an easy consequence of a Hardy-Littlewood Tauberian theorem.

we have

$$E\{Y_1+\ldots+Y_{v_0}|X_0=i\}=m_{ii}\sum_{j=1}^{\infty}\frac{x_j}{m_{jj}}, \tag{Xa}$$ [9]

$$E\{(Y_1+\ldots+Y_{v_0})^2|X_0=i\}$$

$$=m_{ii}\sum_{j=1}^{\infty}\frac{x_j^2}{m_{jj}}+2m_{ii}\sum_{\substack{j=1\\ j\neq i}}^{\infty}\frac{x_j}{m_{jj}}\sum_{k=1}^{\infty}\frac{m_{ji}+m_{ik}-m_{jk}}{m_{kk}}x_k. \tag{Xb}$$

PROOF. It is more convenient to consider new variables Z_n defined as follows:

$$Z_n=\begin{cases}0 \text{ if } X_v=i \text{ for some } v,\ 1\le v<n;\\ x_j \text{ if } X_n=j \text{ and } X_v\neq i,\ 1\le v<n,\end{cases}$$

where j may be i in the last-written line. Evidently we have then

$$E\left\{\sum_{n=1}^{v_0}Y_n\middle|X_0=i\right\}=E\left\{\sum_{n=1}^{\infty}Z_n\middle|X_0=i\right\}$$

$$=\sum_{n=1}^{\infty}E\left\{Z_n\middle|X_0=i\right\}=\sum_{n=1}^{\infty}\sum_{j=1}^{\infty}{}_iP_{ij}^{(n)}x_j$$

$$=\sum_{j=1}^{\infty}{}_iP_{ij}^{*}x_j=\sum_{j=1}^{\infty}\frac{m_{ii}}{m_{jj}}x_j,$$

by (IX) with $i=k$.

Furthermore, we have

$$E\left\{\left(\sum_{n=1}^{v_0}Y_n\right)^2\middle|X_0=i\right\}=E\left\{\left(\sum_{n=1}^{\infty}Z_n\right)^2\middle|X_0=i\right\}$$

$$=E\left\{\sum_{n=1}^{\infty}\left(Z_n^2+2\sum_{1\le r<s\le n}Z_rZ_s\right)\middle|X_0=i\right\}.$$

As before, we obtain readily

$$E\left\{\sum_{n=1}^{\infty}Z_n^2|X_0=i\right\}=\sum_{j=1}^{\infty}\frac{m_{ii}}{m_{jj}}x_j^2.$$

Next we have

$$\sum_{r=1}^{\infty}\sum_{s=r+1}^{\infty}E(Z_rZ_s|X_0=i)=$$

$$\sum_{r=1}^{\infty}\sum_{s=r+1}^{\infty}\sum_{\substack{j=1\\ j\neq i}}^{\infty}\sum_{k=1}^{\infty}{}_iP_{ij}^{(r)}x_j\,{}_iP_{jk}^{(s-r)}x_k=\sum_{\substack{j=1\\ j\neq i}}^{\infty}{}_iP_{ij}^{*}x_j\sum_{k=1}^{\infty}{}_iP_{jk}^{*}x_k.$$

By (IX) this reduces to (Xb).

The two expressions on the left sides of (Xa) and (Xb) play an important role in Doblin's Central Limit Theorem for DMCS. We refer the reader to [2] for details. They are here evaluated in what seems to us more tangible terms.

4. From formulas (VI) and (VII) it follows that

$$m(i,j\cup k)+m(i,j\cap k)=m_{ij}+m_{ik}.$$

This relation is in striking resemblance to a familiar formula in the elementary calculus of probabilities, according to which if A and B are any two events then

$$P(A\cup B)+P(A\cap B)=P(A)+P(B).$$

The generalizations of the last relation to any finite number of events is known as Poincaré's formula (see, for example, [9], p. 61); and we immediately suspect that the same may be true for the mean first passage times. This is indeed so. We define $m(i,j_1\cup\ldots\cup j_s)$ and $m(i,j_1\cap\ldots\cap j_s)$ as the obvious extensions from the case $s=2$. We shall also write $j_1\cap\ldots\cap j_s\div j$ to denote $j_1\cap\ldots\cap j_{r-1}\cap j_{r+1}\cap\ldots\cap j_s$ if $j=j_r$ $(1\le r\le s)$ and $j_1\cap\ldots\cap j_s$ if j is not one of the j_r's.

FORMULA XI. *If $j_1,\ldots,j_s$ are distinct states in a positive class to which i also belongs, then*

$$m(i,j_1\cup\ldots\cup j_s)=\sum_{r=1}^{s}m(i,j_r)-\sum_{1\le r_1<r_2\le s}m(i,j_{r_1}\cap j_{r_2})$$

$$+\sum_{1\le r_1<r_2<r_3\le s}m(i,j_{r_1}\cap j_{r_2}\cap j_{r_3})-\ldots+$$

$$(-1)^{s-1}m(i,j_1\cap\ldots\cap j_s). \tag{XI}$$

PROOF. Put

$$e_n^r=\{w:X_0(w)=i,\ X_v(w)\neq j_1,\ldots,j_s \text{ for } 1\le v<n,\ X_n(w)=j_r\}$$

$$e^r=\bigcup_{n=1}^{\infty}e_n^r.$$

We have, as at the beginning of section 3,

$$m(i,j_{r_1}\cap\cdots\cap j_{r_l})$$

$$=c^{-1}\sum_{r=1}^{s}\sum_{n=1}^{\infty}\int_{e_n^r}\{n+m(j_r,j_{r_1}\cap\ldots\cap j_{r_l}\div j_r)\}P(dw)$$

$$=m(i,j_1\cup\ldots\cup j_s)$$

$$+\sum_{r=1}^{s}c^{-1}P(e^r)m(j_r,j_{r_1}\cap\ldots\cap j_{r_l}\div j_r). \tag{13}$$

Substitute (13) into the right side of (XI) and consider the typical term $c^{-1}P(e^r)m(j_r,j_{r_1}\cap\ldots\cap j_{r_l})$, where r is distinct from $r_1,\ldots,r_l$. It appears on the right side of (XI) once in $m(i,j_{r_1}\cap\ldots\cap j_{r_l})$ and once in $m(i,j_r\cap j_{r_1}\cap\ldots\cap j_{r_l})$, with opposite signs; and does not appear in any other terms. Hence its net coefficient on the right side of (XI) is zero. It remains only to consider the term $m(i,j_1\cup\ldots\cup j_s)$. This appears once in every term and hence its net coefficient is

$$\binom{s}{1}-\binom{s}{2}+\binom{s}{3}-\ldots+(-1)^{s-1}=1.$$

Therefore (XI) is established.

[9] Doob pointed out to me that this is a case of Wald's equation for Markov chains.

We remark that trivial as this proof is, it does not exactly correspond with the familiar proof of Poincaré's formula, and we do not know if there is any closer relation between the two apparent twins. We also leave possible extensions suggested by the known extensions of Poincaré's formula to the interested reader.

5. We now give another method of computing ${}_kP^*_{ij}$. This method requires the ergodic theorem (11). An interesting byproduct is the following:

THEOREM 5. *If i, j and k (not necessarily distinct) belong to a positive class, then*

$$\sum_{n=1}^{\infty} \{P^{(n)}_{ik} - P^{(n)}_{jk}\} = \frac{m_{jk} - m_{ik}}{m_{kk}}.$$

PROOF. Using the familiar formula (cf. the remark after (2))

$$P^{(n)}_{ik} = \sum_{v=1}^{n} F^{(v)}_{ik} P^{(n-v)}_{kk},$$

we have

$$\sum_{n=1}^{N} \{P^{(n)}_{ik} - P^{(n)}_{jk}\} = \sum_{v=1}^{N} \{F^{(v)}_{ik} - F^{(v)}_{jk}\} \sum_{n=0}^{N-v} P^{(n)}_{kk}. \quad (14)$$

Substituting from (11), we see that the right side of (14) is, as $N \to \infty$, asymptotically equal to

$$\sum_{v=1}^{N} \{F^{(v)}_{ik} - F^{(v)}_{jk}\} \frac{N-v}{m_{kk}}. \quad (15)$$

Now since

$$F^*_{ik} = F^*_{jk} = 1, \quad \sum_{v=1}^{\infty} vF^{(v)}_{ik} = m_{ik} < \infty, \quad \sum_{v=1}^{\infty} vF^{(v)}_{jk} = m_{jk} < \infty,$$

we have, as $N \to \infty$

$$N \sum_{v=1}^{N} \{F^{(v)}_{ik} - F^{(v)}_{jk}\} = N \sum_{v=N+1}^{\infty} \{F^{(v)}_{ik} - F^{(v)}_{jk}\} \to 0.$$

Using this in (15), we see that its limit as $N \to \infty$ is

$$\lim_{N\to\infty} \frac{1}{m_{kk}} \sum_{v=1}^{N} (-v)\{F^{(v)}_{ik} - F^{(v)}_{jk}\} = \frac{m_{jk} - m_{ik}}{m_{kk}}.$$

We note that theorem 5 gives a convenient determination of the mean first passage times in terms of the transition probabilities; in particular

$$m_{jk} = \left(1 + \sum_{n=1}^{\infty} \{P^{(n)}_{kk} - P^{(n)}_{jk}\}\right) \lim_{n\to\infty} \frac{1}{N} \sum_{n=1}^{N} P^{(n)}_{kk}.$$

We do not know what the situation is in a null class. All we can infer from theorem 1, (IVc), is that if i and j belong to one recurrent class and ${}_jF^*_{ji} \neq {}_iF^*_{ij}$, then

$$\sum_{n=1}^{\infty} \{P^{(n)}_{ii} - P^{(n)}_{ij}\} = \pm\infty.$$

To return to ${}_kP^*_{ij}$. If $j \neq k$, we have evidently

$$\sum_{h\neq k} {}_kP^{(n)}_{ih} P^{(1)}_{hj} = {}_kP^{(n+1)}_{ij}, \quad (16)$$

where, as later, an unspecified summation runs from 1 to ∞. Summing (16) over n, we obtain

$$\sum_{h\neq k} {}_kP^*_{ih} P^{(1)}_{hj} = {}_kP^*_{ij} - {}_kP^{(1)}_{ij},$$

or

$$\sum_{h} {}_kP^*_{ih} P^{(1)}_{hj} = {}_kP^*_{ij} + P^{(1)}_{kj} - P^{(1)}_{ij} \quad (17)$$

since

$${}_kP^*_{ik} = 1, \quad {}_kP^{(1)}_{ij} = P^{(1)}_{ij}.$$

We assert that in general

$$\sum_{h} {}_kP^*_{ih} P^{(n)}_{hj} = {}_kP^*_{ij} + \sum_{v=1}^{n} \{P^{(v)}_{kj} - P^{(v)}_{ij}\}.$$

This is readily shown by induction on n, starting with (17). Now sum from $n=1$ to $n=N$, divide by N, and let $N \to \infty$. By (11) and theorem 5 we obtain

$$\left(\sum_h {}_kP^*_{ih}\right)\frac{1}{m_{jj}} = {}_kP^*_{ij} + \sum_{v=1}^{\infty} \{P^{(v)}_{kj} - P^{(v)}_{ij}\}$$

$$= {}_kP^*_{ij} + \frac{m_{ij} - m_{kj}}{m_{jj}}. \quad (18)$$

Now,

$$\sum_h {}_kP^*_{ih} = \sum_{h=1}^{\infty} \sum_{n=1}^{\infty} {}_kP^{(n)}_{ih}$$

$$= \sum_{n=1}^{\infty} P(X_v \neq k, 1 \leq v < n \mid X_0 = i) = \sum_{n=1}^{\infty} \sum_{v=n}^{\infty} F^{(v)}_{ik} = m_{ik}. \quad (19)$$

(18) and (19) give (IX).

References

[1] A. N. Kolmogorov, Anfangsgründe der Theorie der Markoffschen Ketten mit unendlich vielen möglichen Zuständen, Matematiceskii Sbornik, N. S., vol. **1**, 607 to 610 (1936); Bul. Univ. Etat Moscow **1**, (1937) (in Russian).

[2] W. Doblin, Sur deux problèmes de M. Kolmogoroff concernant les chaines dénombrables, Bul. Soc. Math. France **66**, 210 to 220 (1938).

[3] W. Feller, Fluctuation theory of recurrent events, Trans. Am. Math. Soc. **67**, 98 to 119 (1949). See also [10], chap. 12.

[4] T. E. Harris, First passage and recurrence distributions. Trans. Am. Math. Soc. **73**, 471 to 486 (1952).

[5] K. L. Chung, An ergodic theorem for stationary Markov chains with a countable number of states, Proc. International Congress of Math. **II**, 568 (1950).

[6] K. L. Chung, Lecture notes on Markov chains, Graduate Mathematical Statistics Society, Columbia University, New York, N. Y., (1951).

[7] K. L. Chung and P. Erdös, Probability limit theorems assuming only the first moment I, Mem. Am. Math. Soc. **6** (1951).

[8] K. L. Chung and W. Fuchs, On the distribution of values of sums of random variables, Mem. Am. Math. Soc. **6** (1951).

[9] W. Feller, An introduction to probability theory and its applications (John Wiley & Sons, New York, N. Y., 1950).

[10] P. Lévy, Systèmes Markoviens et stationnaires. Cas dénombrables. Ann. Sci. École Norm. Sup. (3), **68**, 327 to 381 (1952).

WASHINGTON, January 14, 1952.

Reprinted from *Trans. Amer. Math. Soc.* **76** (1954), pp. 397–419
with permission from the American Mathematical Society.

CONTRIBUTIONS TO THE THEORY OF MARKOV CHAINS. II

BY
K. L. CHUNG(1)

Introduction. This is a sequel to my paper [1]. The present developments are largely independent of the previous results except in so far as given in the Appendix. Theorem 1 shows a kind of solidarity among the states of a recurrent class; it generalizes a classical result due to Kolmogorov and permits a classification of recurrent states and classes. In §2 some relations involving the mean recurrence and first passage times are given. In §§3–5 sequences of random variables associated in a natural way with a Markov chain are studied. Theorem 2 is a generalized ergodic theorem which applies to any recurrent class, positive or null. It turns out that in a null class there is a set of numbers which plays the role of stationary absolute probabilities. In the case of a recurrent random walk with independent, stationary steps these numbers are all equal to one and the result is particularly simple. Theorem 3 shows that the kind of solidarity exhibited in Theorem 1 persists in such a sequence; it leads to the clarification of certain conditions stated by Doblin(2) in connection with his central limit theorem. Using a fundamental idea due to Doblin, the weak and strong laws of large numbers, the central limit theorem, the law of the iterated logarithm, and the limit theorems for the maxima of the associated sequence are proved very simply. Owing to the great simplicity of the method it is the conditions of validity of these limit theorems that should deserve attention. Among other things, we shall show by an example that a certain set of conditions, attributed to Kolmogorov, is in reality *not* sufficient for the validity of the central limit theorem. Furthermore, conditions of validity for the strong limit theorems and the limit theorems for the maxima are obtained by a rather natural strengthening of corresponding conditions for the weak limit theorems. A word about the connection of these conditions with martingale theory closes the paper.

1. The sequence of random variables $\{X_n\}$, $n=0, 1, 2, \cdots$, forms a denumerable Markov chain with stationary transition probabilities. The states will be denoted by the non-negative integers(3) $0, 1, 2, \cdots$. The n-step transition probability from the state i to the state j will be denoted by $P_{ij}^{(n)}$ ($P_{ij}^{(1)}=P_{ij}$). Thus we have

$$P_{ij}^{(n)} = P(X_{m+n} = j \mid X_m = i)$$

Presented to the Society, September 1, 1953; received by the editors April 17, 1953.

(1) This research is done on an Air Research and Development Command Project.

(2) Doeblin.

(3) We note that in [1] the states are the positive integers.

for all integers $m \geqq 0$ for which the conditional probability is defined. In the following we shall simply write $m=0$ in such formulas. The initial probabilities, namely the distribution of X_0, are fixed but arbitrary.

The notations used below are the same as in [1]. In particular

$$
\begin{aligned}
{}_kP_{ij}^{(n)} &= P(X_n = j,\ X_v \neq k,\ 1 \leqq v < n \mid X_0 = i),\\
F_{ij}^{(n)} &= P(X_n = j,\ X_v \neq j,\ 1 \leqq v < n \mid X_0 = i),\\
{}_kF_{ij}^{(n)} &= P(X_n = j,\ X_v \neq j, \neq k,\ 1 \leqq v < n \mid X_0 = i),\\
m_{ij}^{(p)} &= \sum_{n=1}^{\infty} n^p F_{ij}^{(n)},\quad {}_km_{ij}^{(p)} = \sum_{n=1}^{\infty} n^p\, {}_kF_{ij}^{(n)},\\
{}_kP_{ij}^{*} &= \sum_{n=1}^{\infty} {}_kP_{ij}^{(n)},\\
F_{ij}^{*} &= \sum_{n=1}^{\infty} F_{ij}^{(n)},\quad {}_kF_{ij}^{*} = \sum_{n=1}^{\infty} {}_kF_{ij}^{(n)}.
\end{aligned}
$$

We shall confine ourselves to one recurrent class, hence all $F_{ij}^{*}=1$.

THEOREM 1([4]). *Let p be a fixed positive real number. If*

$$
m_{ij}^{(p)} < \infty, \qquad m_{ji}^{(p)} < \infty \tag{1}
$$

for a pair of states i and j (distinct or not) in a recurrent class, then the same is true for every pair of states (distinct or not).

Proof. We have, by the usual arguments, if $i \neq j$:

$$
\begin{aligned}
F_{ii}^{(n)} &= \sum_{v=1}^{n-1} {}_iF_{ij}^{(v)} F_{ji}^{(n-v)} + {}_jF_{ii}^{(n)},\\
F_{ij}^{(n)} &= \sum_{v=1}^{n-1} {}_jF_{ii}^{(v)} F_{ij}^{(n-v)} + {}_iF_{ij}^{(n)}.
\end{aligned}
\tag{2}
$$

Multiplying through by n^p and summing from $n=1$ to $n=\infty$, we obtain

$$
m_{ii}^{(p)} = \sum_{v=1}^{\infty} {}_iF_{ij}^{(v)} \sum_{s=1}^{\infty} (s+v)^p F_{ji}^{(s)} + {}_jm_{ii}^{(p)}, \tag{3}
$$

$$
m_{ij}^{(p)} = \sum_{v=1}^{\infty} {}_jF_{ii}^{(v)} \sum_{s=1}^{\infty} (s+v)^p F_{ij}^{(s)} + {}_im_{ij}^{(p)}. \tag{4}
$$

(i) We shall show that (1) implies that

([4]) *Added in proof.* A partial result, for integral values of p, was proved independently by J. L. Hodges and M. Rosenblatt, *Recurrence-time moments in random walks*, Pacific Journal of Mathematics vol. 3 (1953) pp. 127–136.

$$m_{ii}^{(p)} < \infty. \tag{5}$$

If $i=j$ there is nothing to show. Otherwise from (3) we obtain, since $(s+v)^p \leqq 2^p(s^p+v^p)$,

$$m_{ii}^{(p)} \leqq 2^p \sum_{v=1}^{\infty} {}_iF_{ij}^{(v)}\{m_{ji}^{(p)} + v^p\} + {}_jm_{ii}^{(p)} \leqq 2^p\{m_{ji}^{(p)} + {}_im_{ij}^{(p)}\} + {}_jm_{ii}^{(p)}. \tag{6}$$

From (4) we obtain, since $(s+v)^p \geqq v^p$,

$$m_{ij}^{(p)} \geqq {}_jm_{ii}^{(p)} + {}_im_{ij}^{(p)}. \tag{7}$$

Hence if (1) is true, it follows firstly from (7) that ${}_jm_{ii}^{(p)} < \infty$, ${}_im_{ij}^{(p)} < \infty$, then from (6) that $m_{ii}^{(p)} < \infty$. By symmetry we have also $m_{jj}^{(p)} < \infty$.

(ii) We shall show that (5) implies (1) *for all* j(5). From (3) we obtain, since $(s+v)^p \geqq \max\{s^p, v^p\}$,

$$m_{ii}^{(p)} \geqq \max\{{}_iF_{ij}^{*} m_{ji}^{(p)}, {}_im_{ij}^{(p)}\}.$$

Since ${}_iF_{ij}^{*} > 0$ we have $m_{ji}^{(p)} < \infty$; also ${}_im_{ij}^{(p)} < \infty$.

Next we have, from (2),

$$\sum_{n=1}^{N} n^p F_{ij}^{(n)} = \sum_{v=1}^{N-1} {}_jF_{ii}^{(v)} \sum_{s=1}^{N-v} s^p F_{ij}^{(s)} + \sum_{v=1}^{N-1} {}_jF_{ii}^{(v)} \sum_{s=1}^{N-v} \{(s+v)^p - s^p\} F_{ij}^{(s)} + \sum_{n=1}^{N} n^p\, {}_iF_{ij}^{(n)}.$$

If $0<p\leqq 1$, then $(s+v)^p - s^p \leqq v^p$. Since ${}_jm_{ii}^{(p)} \leqq m_{ii}^{(p)} < \infty$, and we have just seen that ${}_im_{ij}^{(p)} < \infty$, it follows on letting $N\to\infty$ that

$$(1 - {}_jF_{ii}^{*})m_{ij}^{(p)} \leqq {}_jm_{ii}^{(p)} + {}_im_{ij}^{(p)} < \infty.$$

Since ${}_jF_{ii}^{*} < 1$, we have $m_{ij}^{(p)} < \infty$.

If $p>1$, we use the inequality

$$(s+v)^p - s^p \leqq pv(s+v)^{p-1} \leqq p2^{p-1}v(s^{p-1} + v^{p-1}).$$

We thus obtain, similarly to the above,

$$(1 - {}_jF_{ii}^{*})m_{ij}^{(p)} \leqq p2^{p-1}\{{}_jm_{ii}^{(1)} m_{ij}^{(p-1)} + {}_jm_{ii}^{(p)}\} + {}_im_{ij}^{(p)}. \tag{8}$$

Now we are going to use induction on $[p]$. Suppose that we have proved that $m_{ii}^{(p-1)} < \infty$ implies $m_{ij}^{(p-1)} < \infty$ for every j. If $m_{ii}^{(p)} < \infty$ then $m_{ii}^{(p-1)} \leqq m_{ii}^{(p)} < \infty$ and so by the induction hypothesis we have $m_{ij}^{(p-1)} < \infty$. Since $p>1$ the assumption $m_{ii}^{(p)} < \infty$ also implies that ${}_jm_{ii}^{(1)} < \infty$. Hence it follows from (8) that $m_{ij}^{(p)} < \infty$.

(5) Henceforth j is a general index, not necessarily the specific one in (1).

Since we have proved that the implication "$m_{ii}^{(p)} < \infty$ implies $m_{ij}^{(p)} < \infty$" is true for all p with $0 < p \leqq 1$, it is now true for all $p > 0$, by induction on $[p]$.

(iii) We have thus far proved that if (1) is true for a pair of i and j then it is also true for *the same i but all j*. Hence by (i) also $m_{jj}^{(p)} < \infty$ for all j.

REMARK 1. The two conditions in (1) cannot be replaced by one of them. It is possible that $m_{ij} < \infty$ but $m_{ji} = \infty$. Consider for example a random walk with a reflecting barrier: $P_{01} = 1$, $P_{n,n+1} = P_{n,n-1} = 1/2$ for $n \geqq 1$. This example also shows that some m_{ij} may be finite in a null class.

REMARK 2. Let $\phi(n)$ be a positive, nondecreasing function of n satisfying $\phi(n+n') \leqq \phi(n) + A\phi(n')$ for some constant $A > 0$ and every pair of positive integers n and n' where A is a positive constant. Define a generalized moment(6) as follows:

$$m_{ij}^{(\phi)} = \sum_{n=1}^{\infty} \phi(n) F_{ij}^{(n)}.$$

Then the theorem remains true if we replace $m^{(p)}$ by $m^{(\phi)}$. This is a slight generalization of the theorem for the case $0 < p \leqq 1$. The proof is exactly the same as in that case.

REMARK 3. For p a positive integer, we have

$$m_{ii}^{(p)} = \sum_{q=0}^{p} \binom{p}{q} {}_{i}m_{ij}^{(p-q)} m_{ji}^{(q)} + {}_{j}m_{ii}^{(p)}, \tag{9}$$

$$ {}_{i}F_{ij}^{*} m_{ij}^{(p)} = \sum_{q=0}^{p-1} \binom{p}{q} {}_{j}m_{ii}^{(p-q)} m_{ij}^{(q)} + {}_{i}m_{ij}^{(p)}. \tag{10}$$

Kolmogorov [12] proved that if $m_{ii} < \infty$ for some i in a recurrent class, then $m_{jk} < \infty$ for all j and k (distinct or not) in the same class. This is part of the assertion of Theorem 1 for $p = 1$. For $p = 2$ the result has an important application to Doblin's central limit theorem; see §4.

Theorem 1 makes it possible to classify recurrent states and classes. One possibility is as follows. We define the "order" of a recurrent state to be the supremum of all numbers $p \geqq 0$ for which $m_{ii}^{(p)} < \infty$. It follows from Theorem 1 that all the states in one recurrent class have the same order, which may therefore also be called the order of the class. Obviously, a positive state or class (in the established terminology(7)) is of order $\geqq 1$, and a null state or class is of order $\leqq 1$. A state or class of order 1 may be positive or null, as will appear shortly. We state the following existence theorem.

For any given $p \geqq 0$ there exists a Markov chain such that all of its states form a class of order p.

(6) I am indebted to J. L. Doob for the suggestion of a generalized moment; unfortunately I do not see how to extend the theorem significantly for all p.

(7) Lévy [15] uses "strongly ergodic" for "positive" and "weakly ergodic" for "null."

This follows from a simple lemma(8) given by Yosida and Kakutani [16], which asserts: *Given a sequence of numbers* f_n, $n \geqq 1$, *such that* $f_n \geqq 0$, $\sum f_n = 1$, *there exists a Markov chain for which* $F_{00}^{(n)} = f_n$. To prove the existence theorem above it suffices to take in the lemma

$$f_n = \frac{C}{n^{p+1}} \quad \text{or} \quad \frac{C'}{n^{p+1} \lg^2 n}$$

where $p > 0$ and C and C' are constants such that $\sum_n f_n = 1$. In either case the state 0 is of order p; however $m_{00}^{(p)} = \infty$ with the first choice, while $m_{00}^{(p)} < \infty$ with the second. For $p = 0$ only the second choice is possible.

2. In this section we consider a positive class and we investigate some relations involving the m_{ij}'s and also give numerical examples.

The following interesting relation is due to T. E. Harris [9]:

$$\lim_{j \to \infty} {}_iF_{ij}^{*} m_{ij} = m_{ii}. \tag{11}$$

We give a new proof here to put it into closer relationship with our developments. Putting $p = 1$ in (9) and (10), we have

$$m_{ii} = {}_iF_{ij}^{*} m_{ji} + {}_im_{ij} + {}_jm_{ii}, \tag{9'}$$

$${}_iF_{ij}^{*} m_{ij} = {}_jm_{ii} + {}_im_{ij}. \tag{10'}$$

On comparing the two formulas we obtain, firstly,

$$m_{ii} = {}_iF_{ij}^{*}(m_{ij} + m_{ji}) \tag{12}$$

also due to Harris ([9]; see (VIIIa) of [1]). Secondly, since $0 \leqq F_{ii}^{(n)} - {}_jF_{ii}^{(n)} \to 0$ as $j \to \infty$, the infinite series for ${}_jm_{ii}$ converges uniformly in j, and

$$\lim_{j \to \infty} {}_jm_{ii} = m_{ii}, \tag{13}$$

it follows from (9') that

$$\lim_{j \to \infty} {}_im_{ij} = 0, \qquad \lim_{j \to \infty} {}_iF_{ij}^{*} m_{ji} = 0.$$

The last relation and (12) imply (11).

As a simple consequence, *we have for every* i

$$\sum_j \frac{1}{m_{ij}} < \infty.$$

This follows from the obvious relation $m_{jj} \leqq m_{ij} + m_{ji}$ and Kolmogorov's

(8) It can be easily proved by the scheme in Example 1 below.

result $\sum_{j=0}^{\infty} 1/m_{jj}=1$.

For further relations we note the following remarkable formula which expresses the second moment in terms of the first moments:

$$m_{ii}^{(2)} = m_{ii}\left(2\sum_{j=0}^{\infty}\frac{m_{ji}}{m_{jj}} - 1\right). \tag{14}$$

To prove (14) we have only to set $f(\cdot)\equiv 1$ in (B) of the Appendix, and use (C) there:

$$\begin{aligned} m_{ii}^{(2)} &= m_{ii} + 2m_{ii}\sum_{j\neq i}\frac{1}{m_{jj}}\sum_{k}\frac{m_{ji}+m_{ik}-m_{jk}}{m_{kk}} \\ &= m_{ii} + 2m_{ii}\sum_{j\neq i}\frac{m_{ji}}{m_{jj}} = m_{ii}\left(2\sum_{j}\frac{m_{ji}}{m_{jj}} - 1\right). \end{aligned}$$

From (14) *it follows that the series*

$$\sum_{j}\frac{m_{ji}}{m_{jj}}$$

converges if and only if $m_{ii}^{(2)}<\infty$. Consequently, in view of Theorem 1, *all such series for different values of i converge or diverge together*. If one side of (14) is infinite, so is the other. In particular if $m_{ii}^{(2)}=\infty$, then

$$\lim_{j\to\infty} m_{ji} = \infty.$$

As a counterpart to the last result we now give an example where $\lim_{j\to\infty} m_{j0}=1$. Define a Markov chain as follows:

$$\begin{aligned} P_{01} &= 1, & P_{12} &= 1, & & \\ P_{n,n+1} &= 1/n^2, & P_{n0} &= 1 - 1/n^2 & & \text{for } n \geqq 2. \end{aligned}$$

By the Borel-Cantelli lemma the state 0 is recurrent; furthermore $m_{00}=2+m_{20}<\infty$ so that all the states form a positive class. Now

$$m_{j0} = 1 - \frac{1}{j^2} + \frac{2}{j^2}\left(1-\frac{1}{(j+1)^2}\right) + \frac{3}{j^2(j+1)^2}\left(1-\frac{1}{(j+2)^2}\right) + \cdots.$$

Hence $\lim_{j\to\infty} m_{j0}=1$.

We now give another example to illustrate the possible asymptotic behavior of the m_{ij}'s. This example will be used in §4 to disprove a statement attributed to Kolmogorov.

EXAMPLE 1. Define a Markov chain as follows: for $n\geqq 0$

$$P_{n0} = 1 - p_n, \qquad P_{n,n+1} = p_n.$$

Put $p_0=1$ and put

$$\pi_0 = 1, \pi_n = p_0 p_1 \cdots p_{n-1}, f_n = \pi_{n-1} - \pi_n.$$

According to the lemma of Yosida and Kakutani, we may take

$$f_0 = 0, \qquad f_n = \frac{C}{n^3 \lg^2 n} \quad \text{for} \quad n \geqq 2, \qquad \sum_{n=1}^{\infty} f_n = 1.$$

Then

$$\pi_n = \sum_{v=n+1}^{\infty} \frac{C}{v^3 \lg^2 v} \sim \frac{C}{2n^2 \lg^2 n}.$$

Since $P_{ij}^{(n)} = P_{i,j-1}^{(n-1)} p_{j-1}$ for $0 \leqq i < j$ and we are obviously in a positive class, we have by Kolmogorov's ergodic theorem,

$$\frac{1}{m_{jj}} = \frac{p_{j-1}}{m_{j-1,j-1}}.$$

Hence

$$\frac{1}{m_{jj}} = \frac{\pi_j}{m_{00}} \sim \frac{C}{2m_{00}} \frac{1}{j^2 \lg^2 j}.$$

On the other hand, by an easy calculation

$$\begin{aligned} m_{j0} &= (1 - p_j) + 2p_j(1 - p_{j+1}) + 3p_j p_{j+1}(1 - p_{j+2}) + \cdots \\ &= \frac{1}{\pi_j} (f_{j+1} + 2f_{j+2} + 3f_{j+3} + \cdots) \sim j. \end{aligned}$$

It follows in particular that

$$\sum_j \frac{m_{j0}}{m_{jj}} < \infty.$$

This checks with a result above since $m_{00}^{(2)} = \sum_{n=1}^{\infty} n^2 f_n < \infty$.

3. In this section we study certain sequences of random variables associated with the Markov chain $\{X_n\}$, $n \geqq 0$. Let $f(\cdot)$ be a real-valued function defined on the non-negative integers and consider the sequence $\{f(X_n)\}$, $n \geqq 0$. If f has a unique inverse, then the new sequence is also a Markov chain; in general this need not be true. A special case which has been frequently discussed (see [14, p. 335] and [8, p. 342]) is the case in which $f(\cdot) = \delta_{i\cdot}$ for a certain i, where δ is the Kronecker symbol. We shall investigate the asymptotic properties of the partial sums

$$S_n = \sum_{v=0}^{n} f(X_v)$$

as $n \to \infty$ and prove some of the classical limit theorems for them. In the

special case just mentioned these results are easily obtained and one might hope to extend them to the general case by a "linear extension." Such an approach however has not been carried out. Instead we shall use an idea due to Doblin [3](9), as follows:

Let i be a given state and let $0 \leqq v_1 < v_2 < \cdots$ be the successive values of v for which $X_v = i$. Let

$$Y_s = \sum_{v=v_s+1}^{v_{s+1}} f(X_v).$$

The Y_s, $s \geqq 1$, are independent random variables with a common distribution, and we have

$$S_n = \sum_{v=0}^{v_1} f(X_v) + \sum_{s=1}^{l-1} Y_s + \sum_{v=v_l+1}^{n} f(X_v). \tag{15}$$

Thus the asymptotic properties of S_n are closely related to those of $\sum_{s=1}^{l-1} Y_s$, whose behavior is classic. It remains to explore this relation at greater length.

We begin with a strong limit theorem which applies to any recurrent class. We state the following lemma.

LEMMA. *Write* ${}_iP_{ij}^* = E_{ij}$. *Then for every* i, j, *and* k *we have*

$$E_{ij}E_{jk} = E_{ik}.$$

In particular $E_{ii} = 1$ *and* $E_{ij}E_{ji} = 1$.

Proof. According to (D) of the Appendix we have

$$E_{ij} = \lim_{n\to\infty} \frac{\sum_{v=0}^{n} P_{jj}^{(v)}}{\sum_{j=0}^{n} P_{ii}^{(v)}} \tag{16}$$

where the limit is finite and not zero. The lemma follows at once.

For a given function f we set

$$I(f) = \sum_{j=0}^{\infty} E_{ij} f(j).$$

It follows from the lemma that the finiteness of $I(|f|)$ is independent of the choice of i. Furthermore, if g is another function and if $I(|f|) < \infty$, $I(|g|) < \infty$, $I(g) \neq 0$, then the number $I(f)/I(g)$ is independent of i.

(9) Here is an instance of an idea, so simple and akin to a familiar one (that of a recurrent event), and yet so new. Doblin himself seemed to have taken some time discovering it; Kolmogorov in [13] gave prominence to it.

THEOREM 2(10). *If* $I(|f|)<\infty$, $I(|g|)<\infty$ *and* $I(g)\neq 0$, *then*

$$P\left\{\lim_{n\to\infty}\frac{\sum_{v=0}^{n} f(X_v)}{\sum_{v=0}^{n} g(X_v)}=\frac{I(f)}{I(g)}\right\}=1.$$

Proof. If $I(|f|)<\infty$, we have evidently (cf. the proof of (A), Appendix),

$$E(Y_1)=\sum_{n=1}^{\infty}\sum_{j=0}^{\infty} {}_iP_{ij}^{(n)}f(j)=\sum_{j=0}^{\infty} {}_iP_{ij}^{*}f(j)=I(f). \tag{17}$$

Let $l=l(n)$ be defined by $v_{l(n)}\leqq n<v_{l(n)+1}$. Suppose first that $f(\cdot)>0$. Let $Y'=\sum_{v=0}^{v_1} f(X_v)$. Since v_1 is finite with probability one, regardless of the initial probabilities, so is Y'. We have from (15)

$$\sum_{s=1}^{l-1} Y_s\leqq\sum_{v=0}^{n} f(X_v)\leqq Y'+\sum_{s=1}^{l} Y_s. \tag{18}$$

Since the state i is recurrent we have $\lim_{n\to\infty} l(n)=\infty$ with probability one. By the strong law of large numbers of Kolmogorov and Khintchine, we have

$$P\left\{\lim_{n\to\infty}\frac{1}{l(n)}\sum_{s=1}^{l(n)} Y_s=E(Y_1)\right\}=1.$$

It follows from this and (18) that

$$P\left\{\lim_{n\to\infty}\frac{1}{l(n)}\sum_{v=0}^{n} f(X_v)=I(f)\right\}=1. \tag{19}$$

The same result holds if $f(\cdot)<0$. Hence it holds in general if $I(|f|)<\infty$.

Replacing f by g in (19) and combining the two results we obtain the theorem.

Theorem 2 may be regarded as an ergodic theorem in the form given by E. Hopf [11, p. 47]. An interesting feature is its validity even though there may be no stationary absolute probabilities for the Markov chain(11). Such probabilities are given by the ergodic limits $P_j=\lim_{n\to\infty} d^{-1}P_{jj}^{(nd)}$, $j\geqq 0$, where d is the period of the class, provided that they are all positive—in other words, if we are in a positive class. Theorem 2 asserts that in any recurrent class, positive or null, we may take instead of the P_j the numbers $E_{ij}={}_iP_{ij}^{*}$, $j\geqq 0$, for any choice of i. Note that $\sum_{j=0}^{\infty} E_{ij}$ converges if and only if the class

(10) After this paper was finished I was informed that T. E. Harris and H. E. Robbins [17] obtained a more general formulation by means of ergodic theory. In our (discrete space) case metric transitivity can also be proved by martingale theory, as pointed out by J. L. Marty. The present proof remains the simplest in this special case.

(11) For an elucidation of the matter see [17].

is positive and then $E_{ij}=m_{ii}/m_{jj}$ (see (7) of the Appendix). In this case the theorem specializes to Theorem 6 below, and becomes also a consequence of Birkhoff's theorem (metric transitivity can be proved by Theorem 1.1, p. 460 of [6]). These numbers E_{ij} possess another property of the stationary absolute probabilities P_j, namely, they satisfy the familiar system of equations

$$u_k = \sum_{j=0}^{\infty} u_j P_{jk}. \tag{20}$$

The proof(12) is immediate since (noting that ${}_iP_{ii}^n = F_{ii}^n$)

$$\sum_{j=0}^{\infty} {}_iP_{ij}^{*} P_{jk} = \sum_{n=1}^{\infty}\sum_{j=0}^{\infty} {}_iP_{ij}^{(n)} P_{jk} = \sum_{n=1}^{\infty} \left\{ {}_iP_{ik}^{(n+1)} + {}_iP_{ii}^{(n)} P_{ik} \right\}$$
$$= \sum_{n=2}^{\infty} {}_iP_{ik}^{(n)} + P_{ik} = {}_iP_{ik}^{*}.$$

In fact, Derman also proved that the E_{ij}, except for a constant of proportionality, are the only solutions of the system (20) which are all positive.

COROLLARY 1. *Let $N_j(n)$ and $N_k(n)$ be the number of v's, $0 \leqq v \leqq n$, for which $X_v = j$ and $= k$ respectively, then*

$$P\left\{ \lim_{n\to\infty} \frac{N_j(n)}{N_k(n)} = \frac{E_{ij}}{E_{ik}} \right\} = 1.$$

This was proved by Harris [9] and Lévy [14], independently. It is the special case of Theorem 2 with $f(\cdot) = \delta_{j\cdot}$ and $g(\cdot) = \delta_{k\cdot}$, where δ is the Kronecker symbol.

COROLLARY 2. *Suppose $X_n = X_0 + \xi_1 + \cdots + \xi_n$, $n \geqq 0$, where X_0 and the ξ's are independent, integral-valued random variables and where the ξ's have a common distribution such that every integer is a recurrent value for the sequence $\{X_n\}$, $n \geqq 0$; the last condition is satisfied if $E(\xi_1) = 0$ (see [2]). If f and g are real-valued functions on the integers such that*

$$\sum_{j=-\infty}^{\infty} |f(j)| < \infty, \qquad \sum_{j=-\infty}^{\infty} |g(j)| < \infty, \qquad \sum_{j=-\infty}^{\infty} g(j) \neq 0,$$

then

$$P\left\{ \lim_{n\to\infty} \frac{\sum_{v=1}^{n} f(X_v)}{\sum_{v=1}^{n} g(X_v)} = \frac{\sum_{j=-\infty}^{\infty} f(j)}{\sum_{j=-\infty}^{\infty} g(j)} \right\} = 1.$$

To prove the corollary we need only note that this is the special case

(12) A longer proof was first given by C. Derman, in his doctoral thesis (Columbia 1954).

where P_{ij} is a function of $i-j$ and where all $E_{ij}=1$. (Here the states are all the integers instead of the non-negative ones; only a notational change is called for.) This fact, remarkable in its probability interpretation(13), follows at once from (16) since $P_{ii}^{(v)}=P_{jj}^{(v)}$ for all i, j and v.

4. We now consider a positive recurrent class and use the same notations as in §3. If the mean of Y_s exists(14) we denote it by

$$E(Y_s) = \mu_i.$$

In particular if we take $f(\cdot)\equiv 1$ then Y_s reduces to the sth recurrence time for the state i. We denote this by

$$T_s = v_{s+1} - v_s, \qquad s \geqq 1,$$

and we have $E(T_s)=m_{ii}$. Let

$$Z_s = Y_s - (\mu_i/m_{ii})\, T_s.$$

The Z_s's are independent random variables with a common distribution whose mean is

$$E(Z_s) = \mu_i - \frac{\mu_i}{m_{ii}} m_{ii} = 0,$$

provided that μ_i exists. We note that

$$Z_s = \sum_{v=v_s+1}^{v_s+1} g(X_v)$$

where

$$g(\cdot) = f(\cdot) - \mu_i/m_{ii}\cdot$$

The variance of Z_s, finite or not, is denoted by

$$0 \leqq E(Z_s^2) = \sigma_i^2 \leqq \infty.$$

It should be noted that σ_i^2 is *not* the variance of Y_s.

We first prove a preliminary result which may be regarded as an extension of Theorem 1.

THEOREM 3. *Either all μ_i, $i\geqq 0$, exist or none exists. Either all σ_i^2, $i\geqq 0$, are finite or they are all infinite*(15).

Proof. Let the v_s be defined as above. Let $j\neq 1$ and let r be the smallest

(13) It states that in any recurrent random walk (with independent, stationary steps) on the integers, the expected number of stops at j between two consecutive stops at i is equal to one for all i and j.

(14) A moment exists if it exists and is finite.

(15) Theorem 3 can be extended to absolute moments of all orders $p>0$.

$n > v_1$ such that $X_n = j$, τ' the smallest $n > \tau$ such that $X_n = j$, and v_{N+1} be the smallest $n > \tau'$ such that $X_n = i$.

Furthermore, let N' and N'' be two random variables defined as follows: N' is the smallest $s \geqq 1$ such that $X_n = j$ for at least one n, $v_s < n \leqq v_{s+1}$; and N'' is the smallest $s \geqq 1$ such that $X_n = j$ for at least one n, $v_{N'+s} < n \leqq v_{N'+s+1}$. Then N' and N'' have the same distribution and $E(N') = \sum_{n=1}^{\infty} n({}_jF_{ii}^*)^{n-1}{}_iF_{ij}^* = 1/{}_iF_{ij}^* < \infty$.

It is clear that $N \leqq N' + N''$ so that $E(N) \leqq 2E(N') < \infty$.

Consider the sum(16)

$$W = Z_1 + \cdots + Z_N.$$

We have

$$E(|W|) \leqq E\left(\sum_{s=1}^{N} |Z_s|\right) \leqq E\left(\sum_{s=1}^{N'+N''} |Z_s|\right) = 2E\left(\sum_{s=1}^{N'} |Z_s|\right).$$

If μ_i exists, each $E(|Z_s|) < \infty$; hence also the conditional expectations

$$E_1 = E\{ |Z_s| \mid X_n \neq j \text{ for all } n,\ v_s < n \leqq v_{s+1}\},$$
$$E_2 = E\{ |Z_s| \mid X_n = j \text{ for at least one } n,\ v_s < n \leqq v_{s+1}\}$$

are both finite. It follows that

$$\begin{aligned} E(|W|) &\leqq 2\left(\sum_{s=1}^{N'} |Z_s|\right) = 2\sum_{n=1}^{\infty} P(N' = n) E\left\{\sum_{s=1}^{N'} |Z_s| \,\Big|\, N' = n\right\} \\ &= 2\sum_{n=1}^{\infty} P(N' = n)\{(n-1)E_1 + E_2\} \\ &= 2(E(N') - 1)E_1 + 2E_2 < \infty. \end{aligned} \tag{21}$$

On the other hand,

$$\begin{aligned} W &= \sum_{n=v_1+1}^{\tau} g(X_n) + \sum_{n=\tau+1}^{\tau'} g(X_n) + \sum_{n=\tau'+1}^{v_{N+1}} g(X_n) \\ &= W_1 + W_2 + W_3, \text{ say.} \end{aligned}$$

The three random variables W_1, W_2, and W_3 are *independent*. This is a slight generalization of Doblin's idea mentioned at the beginning of §3. A formal verification lies in the observation that the "elementary" probabilities concerned:

$$\begin{array}{lll} v_1 = a; & \tau = t; & X_n = i_n \text{ for } a + 1 \leqq n \leqq t; \\ \tau = t; & \tau' = t'; & X_n = i_n \text{ for } t + 1 \leqq n \leqq t'; \\ \tau' = t'; & v_{N+1} = b; & X_n = i_n \text{ for } t' + 1 \leqq n \leqq b \end{array}$$

(16) This idea is suggested by J. Wolfowitz in connection with another proof of Theorem 1.

where $a<t<t'<b$, are in fact multiplicative. Furthermore $E(W_1)=E(W_2)=E(W_3)=0$. Hence applying an inequality due to Doob [5, Lemma 2], we have by (21)

$$E(|W_2|) \leqq E(|W|) < \infty.$$

By the definition of τ and τ' it is clear that $\mu_j=E(W_2)+(\mu_i/m_{ii})m_{jj}$. Since j is arbitrary we have thus proved that the existence of any μ_i implies that of all μ_i.

Similarly, by considering the variances and using the additivity of the variances of independent random variables we see that the finiteness of any σ_i^2 implies that of all σ_i^2.

REMARK. If we set $f(\cdot)\equiv 1$ in Theorem 3, the results reduce to part of the assertion of Theorem 1 for $p=1$ and $p=2$.

We are now in a position to prove the classical limit theorems for the sequence $\{f(X_n)\}$. Note that, thanks to Theorem 3, the hypotheses in the following theorems are invariant if the state i is replaced by any other state. In the following *we shall write* μ_i/m_{ii} *as* M, *and* σ_i^2/m_{ii} *as* B. It will turn out presently that these numbers, if they exist, are in fact independent of the choice of i, thus justifying the notations.

THEOREM 4 (WEAK LAW OF LARGE NUMBERS). *If* $m_{ii}<\infty$ *and* μ_i *exists, then for every* $\epsilon>0$,

$$\lim_{n\to\infty} P\left\{\left|\frac{S_n}{n}-M\right|>\epsilon\right\}=0. \tag{22}$$

Proof. Let the v_ν and T_ν be as before. For every $n\geqq 0$ define $l=l(n)$ by

$$v_l \leqq n < v_{l+1}.$$

We have then

$$S_n-\sum_{\nu=1}^{l-1} MT_\nu = Y'+\sum_{\nu=1}^{l-1} Z_\nu + Y'' \tag{23}$$

where

$$Y'=\sum_{\nu=0}^{v_1} f(X_\nu), \qquad Y''=\sum_{\nu=v_l+1}^{n} f(X_\nu).$$

Since v_1 is finite with probability one, it is clear that $\lim_{u\to\infty} P(|Y'|>u)=0$. However $Y''=Y_n''$ depends on n. To show that $\lim_{u\to\infty} P(|Y_n''|>u)=0$ uniformly with respect to n we need a simple result due to Kolmogorov [12] according to which

$$\lim_{t\to\infty} P(n-v_{l(n)}>t)=0$$

uniformly with respect to n. It follows that

$$P(|Y''| > u) \leqq P\left(\max_{1\leqq s\leqq t}\left|\sum_{v=v_l+1}^{v_l+s} f(X_v)\right| > u\right) + P(n - v_l > t).$$

Hence the left side tends to 0 uniformly in n.

Applying Khintchine's weak law of large numbers to the sequence $\{Z_s\}$, $s\geqq 1$, we obtain

$$\lim_{n\to\infty} P\left(\left|S_n - M\sum_{v=1}^{l-1} T_v\right| > n\epsilon\right) = 0,$$

since $\sum_{v=1}^{l-1} T_v = v_l$ and $P(\lim_{n\to\infty} v_{l(n)}/n = 1) = 1$. This is equivalent to (22).

Corollary. *Under the hypotheses of the theorem all* μ_j/m_{jj}, $0\leqq j<\infty$, *are equal.*

Theorem 5. (Doblin's central limit theorem). *If* $m_{ll}^{(2)} < \infty$ *and* $0<\sigma_l^2 < \infty$ ([17]), *then for every real* x

$$\lim_{n\to\infty} P\left\{\frac{S_n - Mn}{(Bn)^{1/2}} \leqq x\right\} = \frac{1}{(2\pi)^{1/2}}\int_{-\infty}^{x} e^{-y^2/2}dy.$$

The proof of Theorem 5 is completely similar to that of Theorem 4 if we apply the central limit theorem for independent, identically distributed random variables with finite variances. We remark that this proof is considerably simpler in details than the one given originally by Doblin [3], who made several unnecessary estimates. Doblin also proved that if $\sigma_l^2=0$, S_n is the sum of Mn and a constant depending only on the values of X_0 and X_n.

Corollary 1. *Under the hypotheses of the theorem all* σ_j^2/m_{jj}, $0\leqq j<\infty$, *are equal.*

Corollary 2. *If* $f(\cdot)$ *is a bounded function, then Theorem 4 holds under the sole assumption that* $m_{ll}<\infty$, *and Theorem 5 holds under the sole assumption that* $m_{ll}^{(2)}<\infty$ *if we allow a degenerate normal distribution in case* $\sigma_l^2=0$. *In particular, the theorems hold if the number of states is finite without any assumption whatsoever (in a class of mutually communicating states).*

Corollary 3. *The constants* M *and* B *in Theorems 4 and 5 are given by*

$$M = \sum_{j=0}^{\infty} \frac{f(j)}{m_{jj}}, \tag{24}$$

$$B = \sum_{j=0}^{\infty} \frac{g^2(j)}{m_{jj}} + 2\sum_{j=0, j\neq i}^{\infty} \frac{g(j)}{m_{jj}}\sum_{k=0}^{\infty}\frac{m_{ji}+m_{ik}-m_{jk}}{m_{kk}}\, g(k) \tag{25}$$

([17]) Under the assumption $m_{ll}^{(2)}<\infty$ the condition $\sigma_l^2<\infty$ is obviously equivalent to $E(Y_1^2)<\infty$.

provided that the series converge absolutely.

This follows from (A) and (B) of the Appendix. It is easy to see that M may exist without the series on the right side of (24) being absolutely convergent. We give the following example which will be used also in §5.

EXAMPLE 2. Define a Markov chain[18] as follows:

$$P_{0,2i-1} = \frac{6}{\pi^2 i^2}, \qquad P_{2i-1,2} = 1, \qquad P_{2i,0} = 1, \qquad i \geqq 1.$$

Let $f(0)=0$, $f(2i-1)=i$, $f(2i)=-i$ for $i \geqq 1$. Obviously $m_{00}=3$ and $m_{00}^{(2)}=9$. It is easy to see that $m_{ii}=\pi^2 i^2/2$ for $i \geqq 1$. Hence $\sum_j |f(j)|/m_{jj}=\infty$. On the other hand, both M and B are equal to zero.

According to Doblin [3], Kolmogorov stated in a letter to Fréchet in 1937 that the central limit theorem applies to S_n under the following conditions:

$$(26) \qquad m_{ii}^{(2)} < \infty, \qquad \sum_{j=0}^{\infty} \frac{f^2(j)}{m_{jj}} < \infty.$$

Apart from this indirect reference this statement has never been published to our knowledge[19]. We think that it is false, as will be shown by the following example[20].

EXAMPLE 3. Consider Example 1 and set

$$f(j) = m_{j0}^{1/2}, \qquad i \geqq 0.$$

Then $m_{00}^{(2)} < \infty$ as shown there and

$$\sum_j \frac{f^2(j)}{m_{jj}} = \sum \frac{m_{j0}}{m_{jj}} < \infty,$$

so that (26) is satisfied.

We now rewrite (23) in the form

$$(23') \qquad S_n = Y' + \sum_{v=1}^{l-1} Y_v + Y''.$$

It is clear that if the central limit theorem holds for S_n, it would also hold for the sequence of (partial sums of) the independent, identically distributed

---

(18) This chain is periodic with period 3. A trivial modification makes it nonperiodic.

(19) In a subsequent paper [13] on the local central limit theorem Kolmogorov treated only the case of a finite number of states and used Doblin's method.

(20) *Added in proof.* Even if the second condition in (26) is strengthened to $\sum f^r(j)/m_{jj} < \infty$ for an $r > 2$ it is still not sufficient. This can be shown by setting $f(j) = j^{1/r}$ and $f_n = Cn^{-3}(\lg n)^{-2}$ in Example 3. Cf. a theorem of Doblin-Doob [6, Theorem 7.5, p. 228].

random variables Y_s, $s \geqq 1$. Write $Y_1 = Y$. By the construction of the chain in Example 1 we have

$$P(Y = m_{00}^{1/2} + m_{10}^{1/2} + \cdots + m_{n-1,0}^{1/2}) = f_n = \frac{C}{n^3 \lg^2 n}, \qquad n \geqq 2.$$

Since $m_{j0} \sim j$ it follows that $\sum_{j=0}^{n-1} m_{j0}^{1/2} \sim 2n^{3/2}/3$. As $x \to \infty$, we have

$$P(Y > x) \sim \sum_{2n^{3/2}/3 > x} f_n = \sum_{n > (3x/2)^{2/3}} \frac{C}{n^3 \lg^2 n} \sim \frac{C_1}{x^{4/3} \lg^2 x}.$$

On the other hand, we have

$$\int_0^x y^2 d_y P(Y \leqq y) \sim \sum_{2n^{3/2}/3 \leqq x} \left(\frac{2}{3} n^{3/2}\right)^2 f_n = C_2 \sum_{n \leqq (3x/2)^{2/3}} \frac{n^3}{n^3 \lg^2 n} \sim C_3 \frac{x^{2/3}}{\lg^2 x}$$

where C_1, C_2, and C_3 are positive constants. Therefore,

$$\lim_{x \to \infty} \frac{x^2 P(Y > x)}{\int_0^x y^2 d_y P(Y \leqq y)} = \frac{C_1}{C_3} > 0.$$

According to the necessary and sufficient condition ([21]) of Feller and Lévy for the identically distributed case [15, p. 113] this proves that the central limit theorem does not hold for the sequence $\{Y_s\}$; hence it does not hold for the sequence S_n. In particular, Doblin's condition $\sigma_i^2 < \infty$ is not satisfied; in fact as shown above

$$\int_0^\infty y^2 d_y P(Y \leqq y) = \infty.$$

On the other hand, Example 2 shows that Doblin's conditions $m_{ii}^{(2)} < \infty$ and $\sigma_i^2 < \infty$ do not imply (26).

5. The conditions that $m_{ii} < \infty$ and μ_i exists, sufficient for the weak law of large numbers, are not sufficient for the strong law. To see that, we need only return to Example 2. Suppose $X_0 = 0$ with probability one. If $X_v = 0$ for a certain v, then the probability that $f(X_{v+1}) \geqq n$ is equal to

$$\frac{6}{\pi^2} \sum_{i=n}^{\infty} \frac{1}{i^2} \sim \frac{6}{\pi^2 n}.$$

Since $S_{3n} = 0$ for every $n \geqq 0$, we have then $S_{3n+1} \geqq n$ with the above probability. It follows by the Borel-Cantelli lemma that $S_{3n+1} \geqq n$ infinitely often with

([21]) Clearly we can state necessary and sufficient conditions for some of the limit theorems given here if we use the distribution function of Y. To our mind such statements are absolutely futile.

probability one. Hence the strong law of large numbers does not hold, while the weak law holds trivially.

This example suggests that the sufficient condition given in the next theorem, though not necessary([22]), is probably not very far from the truth.

THEOREM 6 (STRONG LAW OF LARGE NUMBERS). *If* $m_{ii} < \infty$ *and*

$$\sum_{j=0}^{\infty} \frac{|f(j)|}{m_{jj}} < \infty, \tag{27}$$

then

$$P\left\{\lim_{n\to\infty} \frac{S_n}{n} = M\right\} = 1.$$

REMARK. If the chain X_n is stationary, then under (27) the number M is simply $E(f(X_0))$.

Proof. The proof of this theorem is implicit in that of Theorem 2; it is also a corollary of that theorem for $g(\cdot) \equiv 1$. We need only note that since we are now in a positive class,

$$E_{ij} = m_{ii}/m_{jj}$$

and $P(\lim_{n\to\infty} l(n)/n = 1/m_{ii}) = 1$.

The following alternative method seems instructive. Referring back to the proof of Theorem 4, we see that we may apply the strong law of large numbers of Kolmogorov and Khintchine to $\{Z_s\}$. It remains thus to prove that

$$P\left(\lim_{n\to\infty} \frac{Y_n''}{n} = 0\right) = 1.$$

A little reflection shows that this is equivalent to

$$P\left\{\max_{v_l < v' \leq v_{l+1}} \left|\sum_{v=v_l+1}^{v'} f(X_v)\right| > n\epsilon \text{ i.o.}\right\} = 0. \tag{28}$$

Let $U_l = \sum_{v=v_l+1}^{v_{l+1}} |f(X_v)|$, then (27) is equivalent to $E(U_l) < \infty$. Since the U_l, $l \geq 1$, have a common distribution it follows by a well known inequality that

$$\sum_n P(U_{l(n)} > n\epsilon) < \infty.$$

Hence by the Borel-Cantelli lemma,

$$P(U_{l(n)} > n\epsilon \text{ i.o.}) = 0.$$

This obviously implies (28) and proves the theorem.

([22]) A slight modification of Example 2 confirms this.

It is interesting that a similar strengthening of the conditions of Theorem 5, obtained by substituting $|f|$ for f, leads to the law of the iterated logarithm.

THEOREM 7 (LAW OF THE ITERATED LOGARITHM)[23]. *If* $m_{ii}^{(2)} < \infty$, *and if*

$$E\left\{\left[\sum_{v=v_1+1}^{v_2} |f(X_v)|\right]^2\right\} < \infty \tag{29}$$

and $B > 0$, *then*

$$P\left\{\limsup_{n\to\infty} \frac{S_n - Mn}{(Bn \lg \lg n)^{1/2}} = 1\right\} = 1.$$

REMARK. Using the random variables U_l introduced in the proof of Theorem 6, the left-hand side of (29) may be written as $E(U_1^2)$. Using (B) of the Appendix it may also be written as the right-hand side of (B) after substituting $|f|$ for f.

Proof. As before we may ignore Y' in (23). According to the law of the iterated logarithm in the form given by Hartman and Wintner [10], we have

$$P\left\{\limsup_{l\to\infty} \frac{\sum_{v=1}^{l-1} Z_v}{(l\sigma_i^2 \lg \lg l)^{1/2}} = 1\right\} = 1.$$

Since $P(\lim_{n\to\infty} l(n)/n = 1/m_{ii}) = 1$ the limit may be taken as $n\to\infty$ and the denominator may be replaced by $(Bn \lg \lg n)^{1/2}$.

It remains to prove, as in the previous proof, that for every $\epsilon > 0$

$$P\{U_{l(n)} > (n \lg \lg n)^{1/2}\epsilon \text{ i.o.}\} = 0.$$

Now it is easy to see that

$$\sum_n P(U_{l(n)} > n^{1/2}) \leqq CE(U_1^2) < \infty$$

where C is some positive constant. Hence we have as before

$$P(U_{l(n)} > n^{1/2} \text{ i.o.}) = 0.$$

This is more than we need to prove the theorem.

Finally, we prove the limit theorems concerning $\max_{0\leqq v\leqq n} S_v$ and $\max_{0\leqq v\leqq n} |S_v|$ respectively.

THEOREM 8[24]. *Under the same hypotheses as in Theorem 7, we have*

---

[23] For the case of a finite number of states see Doblin [3], where a sharper result is proved by a longer method.

[24] The limit joint distribution of max $(S_v - Mv)$ and min $(S_v - Mv)$ can be similarly derived.

$$\lim_{n\to\infty} P\left\{\max_{0\leq v\leq n} (S_v - Mv) \leq x(Bn)^{1/2}\right\} = \begin{cases} 0 & \text{if } x < 0, \\ \left(\dfrac{2}{\pi}\right)^{1/2} \displaystyle\int_0^x e^{-y^2/2} dy & \text{if } x \geq 0, \end{cases}$$

$$\lim_{n\to\infty} P\left\{\max_{0\leq v\leq n} |S_v - Mv| \leq x(Bn)^{1/2}\right\}$$
$$= \frac{4}{\pi}\sum_{m=0}^{\infty} \frac{(-1)^m}{2m+1} \exp\left\{\frac{-(2m+1)^2\pi^2}{8x^2}\right\} \qquad \text{if } x \geq 0.$$

Proof. We shall be brief now. According to results of Erdös and Kac [17], these limit theorems hold for the sequence of random variables $\{Z_n\}$, $n \geq 1$. Hence it is sufficient to show that for every $\epsilon > 0$

$$\lim_{n\to\infty} P\left\{\max_{1\leq s< l}\ \max_{v_s<t\leq v_{s+1}} \left|\sum_{v=v_s+1}^{t} f(X_v)\right| > n^{1/2}\epsilon\right\} = 0,$$

$$\lim_{n\to\infty} P\left\{\max_{a\leq t\leq v_1} \left|\sum_{v=1}^{t} f(X_v)\right| > n^{1/2}\epsilon\right\} = 0, \tag{30}$$

$$\lim_{n\to\infty} P\left\{\max_{v_l<t\leq n} \left|\sum_{v=v_l+1}^{t} f(X_v)\right| > n^{1/2}\epsilon\right\} = 0.$$

We need only consider the first relation, the other two being implicit in the proof of Theorem 4. The probability written there is clearly not greater than

$$nP\left\{\sum_{v=v_1+1}^{v_2} |f(X_v)| > n^{1/2}\epsilon\right\} = nP(U_1 > n^{1/2}\epsilon)$$
$$\leq n\frac{1}{n\epsilon^2}\int_{U_1>n^{1/2}\epsilon} U_1 dPr = \frac{1}{\epsilon^2}\int_{U_1>n^{1/2}\epsilon} U_1 dPr$$

which clearly tends to zero as $n\to\infty$. This proves the theorem.

We forbear from stating further limit theorems which can be proved by the same method, but add a word about the conditions (27) and (29) which appear in Theorems 6, 7, and 8. We note that they are clearly satisfied if $m_{ii}^{(2)} < \infty$ and if f is a bounded function. In particular, Theorems 6, 7, and 8 hold without any assumption (in a class of mutually communicating states) if the number of states is finite. We note furthermore that if we assume only the weaker condition that μ_i exists (in Theorem 6) or $\sigma_i^2 < \infty$ (in Theorems 7 and 8), then we need an inequality of the following form: for some constant $c > 0$ and every $x > 0$,

$$P\left\{\max_{v_1<t\leq v_2} \left|\sum_{v=v_1+1}^{t} f(X_v)\right| > x\right\} \leq cP\left\{\left|\sum_{v=v_1+1}^{v_2} f(X_v)\right| > x\right\}.$$

It is not known what reasonable process, if any, enjoys this property, but

it is somewhat reminiscent of martingales. At least in one case, namely Theorem 8, the condition (29) may be replaced by the weaker one $\sigma_i^2<\infty$ and the additional assumption that $S_n=\sum_{v=0}^n f(X_v)$, $n\geqq 0$, is a martingale. This means that the following equations are satisfied for all $i\geqq 0$:

$$\sum_{j=0}^{\infty} P_{ij}f(j)=f(i),\qquad \sum_{j=0}^{\infty} P_{ij}\,|\,f(j)\,|<\infty.$$

Then the sequence $V_n=\sum_{v=v_1+1}^{v_1+n} f(X_v)$, $n\geqq 1$, is a martingale and the process obtained from V_n by optional stopping:

$$V_1,\cdots,V_{v_2-v_1}$$

is also a martingale (see [6, Theorem 2.1, p. 300]). The process V_n^2, $1\leqq n\leqq v_2-v_1$, is then a semi-martingale (loc. cit., Theorem 1.1, p. 295) and we have (loc. cit., Theorem 3.2, p. 314)

$$nP\left\{\max_{1\leqq t\leqq v_2-v_1}|\,V_t\,|>n^{1/2}\epsilon\right\}\leqq\frac{1}{\epsilon^2}\int_{\max_{1\leqq t\leqq v_2-v_1}|V_t|>n^{1/2}\epsilon} V_{v_2-v_1}^2\,dPr. \tag{31}$$

The conditions $m_{ii}^{(2)}<\infty$ and $\sigma_i<\infty$ imply that $E(V_{v_2-v_1}^2)<\infty$. Since $\max_{1\leqq t\leqq v_2-v_1}|V_t|$ is finite with probability one, the right side of (31) tends to 0 as $n\to\infty$, and this is sufficient to prove the first relation in (30).

Appendix

For the sake of convenience we state and prove some results from [1] which are used in the present paper.

In the notations used at the beginning of §3, we have (Theorem 3 of [1]):

$$E\left\{\sum_{v=v_1+1}^{v_2} f(X_v)\right\}=m_{ii}\sum_{j=0}^{\infty}\frac{f(j)}{m_{jj}}, \tag{A}$$

$$\begin{aligned}E\left\{\left[\sum_{v=v_1+1}^{v_2} f(X_v)\right]^2\right\}&=m_{ii}\sum_{j=0}^{\infty}\frac{f^2(j)}{m_{jj}}\\&\quad+2m_{ii}\sum_{j=0,\,j\neq i}^{\infty}\frac{f(j)}{m_{jj}}\sum_{k=0}^{\infty}\frac{m_{ji}+m_{ik}-m_{jk}}{m_{kk}}f(k),\end{aligned} \tag{B}$$

provided that the series converge absolutely. Furthermore (see (19) of [1]),

$$\sum_{k=0}^{\infty}\frac{m_{ji}+m_{ik}-m_{jk}}{m_{kk}}=m_{ji}. \tag{C}$$

Proof. If $j\neq k$ we have

$$F_{jk}^{(n)}=\sum_{v=0}^{n-1}{}_kP_{jj}^{(v)}\,{}_jF_{jk}^{(n-v)}$$

where ${}_kP_{jj}^{(0)}=1$. Summing over n we obtain

$$1 = F^*_{jk} = {}_jF^*_{jk}(1 + {}_kP^*_{jj}). \tag{1}$$

Further, if $j \neq k$ we have ${}_kP_{ij}^{(n)} = \sum_{v=1}^{n} {}_kF_{ij}^{(v)}\, {}_kP_{jj}^{(n-v)}$. Summing over n we obtain

$${}_kP^*_{ij} = {}_kF^*_{ij}(1 + {}_kP^*_{jj}). \tag{2}$$

From (1) and (2) we have

$${}_kP^*_{ij} = \frac{{}_kF^*_{ij}}{{}_jF^*_{jk}}. \tag{3}$$

Suppose $j \neq k$. Let v' be the smallest value of $v > v_1$ for which there exist n_1 and n_2 such that $v_1 < n_1 \leqq v'$, $v_1 < n_2 \leqq v'$ and $X_{n_1} = j$, $X_{n_2} = k$. The expectation $E(v' - v_1)$ may be suggestively denoted by $m(i, j \text{ and } k)$. Recalling that ${}_kF^*_{ij}$ is the probability, starting from i, of reaching j before k, it is not difficult to see that

$$m(i, j \text{ and } k) = m_{ij} + {}_kF^*_{ij} m_{jk} = m_{ik} + {}_jF^*_{ik} m_{kj}. \tag{4}$$

Since in a recurrent class ${}_kF^*_{ij} + {}_jF^*_{ik} = 1$ we deduce from (4) that

$$m_{ik} + m_{kj} - m_{ij} = {}_kF^*_{ij}(m_{jk} + m_{kj}). \tag{5}$$

Replacing both i and k by j and j by k in (5), we have

$$m_{jj} = {}_jF^*_{jk}(m_{kj} + m_{jk}). \tag{6}$$

From (3), (5), and (6) we obtain then

$${}_kP^*_{ij} = \frac{m_{ik} + m_{kj} - m_{ij}}{m_{jj}}. \tag{7}$$

This is the main formula.

Summing (7) over j we obtain

$$\begin{aligned}\sum_{j=0}^{\infty} \frac{m_{ik} + m_{kj} - m_{ij}}{m_{jj}} &= \sum_{j=0}^{\infty} {}_kP^*_{ij} = \sum_{j=0}^{\infty}\sum_{n=1}^{\infty} {}_kP_{ij}^{(n)} \\ &= \sum_{n=1}^{\infty}\sum_{v=n}^{\infty} F_{ik}^{(v)} = m_{ik}.\end{aligned}$$

Thus (C) is proved. Note that the m_{jj}'s on both sides cancel out.

To prove (A) and (B) we introduce new random variables X'_n, $n \geqq 1$, as follows

$$X'_n = \begin{cases} 0 & \text{if } X_v = i \text{ for some } v\colon\ 1 < v < n; \\ f(X_n) & \text{if } X_v \neq i \text{ for all } v\colon\ 1 \leqq v < n. \end{cases}$$

We have, if the series involved converge absolutely:

$$E\left\{\sum_{v=v_1+1}^{v_2} f(X_v)\right\} = E\left\{\sum_{n=1}^{\infty} X'_n \mid X_0 = i\right\} = \sum_{n=1}^{\infty} E\{X'_n \mid X_0 = i\}$$
$$= \sum_{n=1}^{\infty}\sum_{j=0}^{\infty} {}_iP_{ij}^{(n)} f(j) = \sum_{j=0}^{\infty} {}_iP^*_{ij} f(j) = \sum_{j=0}^{\infty} \frac{m_{ii}}{m_{jj}} f(j)$$

by (7) with $k=i$. Thus (A) is proved.

Similarly, we have

$$E\left\{\left[\sum_{v=v_1+1}^{v_2} f(X_v)\right]^2\right\} = E\left\{\sum_{n=1}^{\infty}\left[X_n'^2 + 2\sum_{\tau=1}^{\infty}\sum_{s=\tau+1}^{\infty} X'_\tau X'_s\right]\middle| X_0 = i\right\}. \tag{8}$$

As before,

$$E\left\{\sum_{n=1}^{\infty} X_n'^2 \mid X_0 = i\right\} = \sum_{j=0}^{\infty} \frac{m_{ii}}{m_{jj}} f^2(j). \tag{9}$$

Next we have

$$\begin{aligned}\sum_{\tau=1}^{\infty}\sum_{s=\tau+1}^{\infty} E(X'_\tau X'_s \mid X_0 = i) &= \sum_{\tau=1}^{\infty}\sum_{s=\tau+1}^{\infty}\sum_{j=0, j\neq i}^{\infty}\sum_{k=0}^{\infty} {}_iP_{ij}^{(\tau)} f(j)\, {}_iP_{jk}^{(s-\tau)} f(k) \\ &= \sum_{j=0, j\neq i}^{\infty} {}_iP^*_{ij} f(j) \sum_{k=0}^{\infty} {}_iP^*_{jk} f(k).\end{aligned} \tag{10}$$

Substituting (9) and (10) into (8) and using (7) we obtain (B).

In a recurrent class we have (Theorem 1 of [1]):

$$0 < {}_iP^*_{ij} = \lim_{n\to\infty} \frac{\sum_{v=0}^{n} P_{jj}^{(v)}}{\sum_{v=0}^{n} P_{ii}^{(v)}} < \infty. \tag{D}$$

Proof. That $0 < {}_iP^*_{ij} < \infty$ follows from (3). Now $P_{ij}^{(n)} = \sum_{v=0}^{n} P_{ii}^{(v)}\, {}_iP_{ij}^{(n-v)}$ where ${}_iP_{ij}^{(0)} = 0$. Summing over n,

$$\sum_{n=0}^{N} P_{ij}^{(n)} = \sum_{v=0}^{N} P_{ii}^{(v)} \sum_{n=0}^{N-v} {}_iP_{ij}^{(n)}.$$

Since $\lim_{N\to\infty} \sum_{n=0}^{N-v} {}_iP_{ij}^{(n)} = {}_iP^*_{ij}$, it follows easily that

$$\lim_{N\to\infty} \frac{\sum_{n=0}^{N} P_{ij}^{(n)}}{\sum_{v=0}^{N} P_{ii}^{(v)}} = {}_iP^*_{ij}.$$

On the other hand it is trivial that

$$\lim_{N\to\infty} \frac{\sum_{n=0}^{N} P_{ij}^{(n)}}{\sum_{n=0}^{N} P_{ii}^{(n)}} = 1.$$

Thus (D) is proved.

References

1. K. L. Chung, *Contribution to the theory of Markov chains* I, Journal of Research of the National Bureau of Standards vol. 50(1953) pp. 203–208.

2. K. L. Chung and W. H. J. Fuchs, *On the distribution of values of sums of random variables*, Memoirs of the American Mathematical Society, no. 6, 1951, 12 pp.

3. W. Doblin, *Sur deux problèmes de M. Kolmogoroff concernant les chaines dénombrables*, Bull. Soc. Math. France vol. 52 (1938) pp. 210–220.

4. ———, Thèse, Paris (printed in Bucharest), 1938.

5. J. L. Doob, *The law of large numbers for continuous stochastic processes*, Duke Math. J. vol. 6 (1940) pp. 290–306.

6. ———, *Stochastic processes*, New York, Wiley, 1953.

7. P. Erdös and M. Kac, *On certain limit theorems of the theory of probability*, Bull. Amer. Math. Soc. vol. 52 (1946) pp. 292–302.

8. W. Feller, *An introduction to probability theory and its applications*, New York, Wiley, 1950.

9. T. E. Harris, *First passage and recurrence distributions*, Trans. Amer. Math. Soc. vol. 73 (1952) pp. 471–486.

10. P. Hartman and A. Wintner, *On the law of the iterated logarithm*, Amer. J. Math. vol. 63 (1941) pp. 169–176.

11. E. Hopf, *Ergodentheorie*, Ergebnisse der Mathematik, vol. 5, no. 2, 1937.

12. A. N. Kolmogorov, *Anfangsgründe der Theorie der Markoffschen Ketten mit unendlichen vielen möglichen Zuständen*, Mat. Sbornik N.S. vol. 1 (1936) pp. 607–610; Bull. de l'Univ. d'Etat Moscou (in Russian) vol. 1 (1937). For an exposition in English see Chapter 1 of K. L. Chung, *Lecture notes on Markov chains*, Columbia Graduate Mathematical Statistics Society, New York, 1951.

13. ———, *A local limit theorem for classical Markov chains* (in Russian), Izvestiya Akad. Nauk SSSR. Ser. Mat. vol. 13 (1949) pp. 281–300.

14. P. Lévy, *Systèmes markoviens et stationnaires; cas dénombrable*, Ann. École Norm. (3) vol. 68 (1952) pp. 327–381.

15. ———, *Théorie de l'addition des variables aléatoires*, Paris, Gauthier-Villars, 1937.

16. K. Yosida and S. Kakutani, *Markoff process with an enumerable infinite number of states*, Jap. J. Math. vol. 16 (1939) pp. 47–55.

17. T. E. Harris and H. E. Robbins, *Ergodic theory of Markov chains admitting an infinite invariant measure*, Proc. Nat. Acad. Sci. U.S.A. vol. 39 (1953) pp. 860–864.

Cornell University,
Ithaca, N. Y.
Syracuse University,
Syracuse, N. Y.

Reprinted from *Trans. Amer. Math. Soc.* **81** (1956), pp. 195–210
with permission from the American Mathematical Society.

SOME NEW DEVELOPMENTS IN MARKOV CHAINS

BY
K. L. CHUNG(1)

1. **Introduction.** The purpose of this paper is to point out some new connections between the sample function behavior and the analytical properties of the transition probability functions of a continuous parameter Markov chain with stationary transition probabilities. The main idea is that of a *post-exit process* derived from the original process by considering its evolution after the exit from a stable state. This leads to various relations between conditional probabilities all of which are "intuitively obvious" but require sometimes painstaking proofs if probability theory like any other branch of mathematics is to be treated as a discipline in logic. These relations imply certain analytical properties of the transition functions, obtained recently by D. G. Austin by purely analytical means, and exhibit them in connection with other quantities introduced in this paper. They also complete some results due to Doob and Lévy. The well-known differential equations of Kolmogorov are seen to be limiting cases of certain more generally valid differential equations involving a continuous parameter. One of these expresses the fundamental transition property of the post-exit process, and the other a similar property of the renewal density of an imbedded renewal process.

2. **A conditional probability.** So far as possible we use the terminology and notation of [6]. Let $\{x(t), t\geqq 0\}$ be a Markov chain with initial distribution $\{p_i\}$ and stationary transition matrix $\{p_{ij}(t)\}$, $t\geqq 0$, $i, j=1, 2, \cdots$, such that $\lim_{t\to 0} p_{ij}(t)=\delta_{ij}=p_{ij}(0)$. Then the $p_{ij}(\cdot)$ are all uniformly continuous in $[0, \infty)$, since $|p_{ij}(t+h)-p_{ij}(t)|\leqq 1-p_{ii}(h)$. According to Theorem II.2.6 of [6] we may take $\{x(t)\}$ to be separable relative to the class of closed sets and measurable. Moreover the denumerable set satisfying the conditions of the separability definition will be taken to be the set of rational numbers of the form $r2^{-m}$ (see [6, pp. 56–60]). Separability requires the adjunction of the value ∞ to the range of $x(\cdot, w)$. We define this value ∞ as an *adjoined state* with $p_\infty=0$ and $p_{\infty i}(t)=p_{i\infty}(t)=\delta_{i\infty}$ for all $t\geqq 0$. We have then $P\{x(t)=\infty\}=0$ for every $t\geqq 0$. In the sequel, the term "state," unless specified to the contrary, shall exclude the adjoined state. The limit

$$q_i = p'_{ii}(0) = \lim_{t\downarrow 0} \frac{1-p_{ii}(t)}{t}$$

Presented to the Society, December 29, 1955; received by the editors May 11, 1955.
(1) Research connected with an Air Development and Research Command Project at Syracuse University, OSR Contract No. AF(18)600–760.

exists for every i (Theorem 9 of [3]). Following Lévy the state i is called *stable* or *instantaneous* according as q_i as finite or infinite. We refer to [2] for the foundations of the theory of Markov chains under consideration. Although knowledge of these foundations will be necessary for a thorough understanding of what follows, we shall strive to make the present paper readable by itself.

Let i be a stable state with $q_i>0$; such a state always exists unless $P\{x(t) \equiv x(0), 0\leqq t<\infty\}=1$ [8, p. 375]. Suppose that $P\{x(0)=i\}=1$. Let $\lambda=\lambda_i(w)$ be the "first sojourn time" in the state i, namely the length of the first t-interval in which $x(t, w)\equiv i$ (see Theorem 1 of [2]). Then $P\{\lambda\leqq t\}=1-e^{-q_i t}$, $t\geqq 0$ [3, p. 54]. Let j be an arbitrary state (not ∞!) and define

$$\alpha = \alpha_{ij}(w) = \inf_{\lambda(w)<t, x(t,w)=j} t.$$

If $j\neq i$, α is the "first entrance time into j"; if $j=i$, α is the "second entrance time into i" (the first being zero by hypothesis). It is easily shown that α is a random variable in the broad sense, namely a measurable w-function defined on a measurable w-set whose probability may be less than one. We define its distribution function in the broad sense by $F_{ij}(t)=P\{\alpha\leqq t\}$.

Now it can be proved that the two random variables λ and $\alpha-\lambda$ (which may be zero with positive probability) are independent([2]). This is a special case of Theorem 5 of [2], but we give a simple proof as follows. Let us first note that the distribution of $\alpha-\lambda$ may be derived as follows. It can be shown that([3]) $\alpha(w)$ considered as a point on the t-axis is the limit from the right of points of $S_j(w)=\{t: x(t, w)=j\}$; hence we have

$$\begin{aligned} P\{\alpha - \lambda > u \mid x(0) = i\} = \lim_{h\to 0} \sum_{n=1}^{\infty} P\{x(rh) = i, 0 \leqq r < n; x(nh) \neq i, \neq j; \\ x(rh) \neq j, n+1 \leqq r \leqq n+1+[u/h] \mid x(0) = i\} \end{aligned}$$

where $h=2^{-m}$, $m\to\infty$. Now for each $s>0$ define two random variables $\lambda_s=\lambda_s(w)$ and $\alpha_s=\alpha_s(w)$ on the set $\{w: x(s, w)=i\}$ as follows: $\lambda_s(w)$ is the supremum of T such that $x(t, w)\equiv i$, $s\leqq t<s+T$; $\alpha_s(w)$ is the infimum of t such that $t>\lambda_s(w)$ and $x(t, w)=j$. Thus λ_0 and α_0 reduce to the previous λ

([2]) The random variables $z_1(w), \cdots, z_n(w)$, with domains of definition $\Lambda_1, \cdots, \Lambda_n$, are said to be independent iff $P\{\bigcap_{k=1}^n \Lambda_k[z_k(w)\leqq c_k]\}/P\{\bigcap_{k=1}^n \Lambda_k\}=\prod_{k=1}^n P\{\Lambda_k[z_k(w)\leqq c_k]\}/P(\Lambda_k)$ for every real $c_1, \cdots, c_k$.

([3]) In fact, the set $S_j(w)$ is dense in itself (see [2, §4, (ii)]).

and α. It is then clear that we have by the preceding derivation and stationarity

$$P\{\alpha - \lambda > u \mid x(0) = i\} = P\{\alpha_s - \lambda_s > u \mid x(s) = i\}.$$

Now by definition we have for every $s>0$,

$$\begin{aligned} P\{\lambda \geqq s, \alpha - \lambda > u \mid x(0) = i\} \\ &= P\{x(t) \equiv i, 0 < t < s; \alpha - \lambda > u \mid x(0) = i\} \\ &= P\{x(t) \equiv i, 0 < t \leqq s \mid x(0) = i\} P\{\alpha_s - \lambda_s > u \mid x(s) = i\} \\ &= P\{\lambda \geqq s \mid x(0) = i\} P\{\alpha - \lambda > u \mid x(0) = i\} \end{aligned}$$

where we have used the fact that the distribution of λ is continuous. This relation is equivalent to the independence of λ and $\alpha-\lambda$ under the hypothesis $x(0)=i$. Hence if $G_{ij}(\cdot)$ denotes the distribution function in the broad sense of $\alpha-\lambda$, then

$$(1) \qquad P\{\alpha \leqq u \mid \lambda\} = G_{ij}(u - \lambda)$$

for all u with probability one. We now prove that

$$(2) \qquad P\{x(t) = j \mid \lambda, \alpha\} = p_{jj}(t - \alpha)$$

for almost all w on the set $\{\alpha \leqq t\}$. This is a consequence of Theorem 6 of [2] but can be proved directly as follows. If $i \neq j$, we have, taking both h and δ below to be negative integral powers of 2,

$$\begin{aligned} P\{\lambda \leqq s_1; \alpha \leqq u_1; x(t) = j\} \\ &= \lim_{h\to 0} \sum_{1 \leqq n \leqq u_1/h} P\{\lambda \leqq s_1; (n-1)h < \alpha \leqq nh; x(t) = j\} \\ (3) \quad &= \lim_{h\to 0} \sum_{1\leqq n \leqq u_1/h} \lim_{\delta\to 0} \sum_{(n-1)h/\delta < m \leqq nh/\delta} P\{\lambda \leqq s_1; x(r\delta) \neq j, \\ &\qquad (n-1)h/\delta < r < m; x(m\delta) = j\} p_{jj}(t - m\delta) \\ &= \lim_{h\to 0} \sum_{1 \leqq n \leqq u_1/h} P\{\lambda \leqq s_1; (n-1)h < \alpha \leqq nh\} p_{jj}(t - nh + \theta h) \\ &\qquad\qquad \text{where } 0 \leqq \theta \leqq 1, \\ &= \int_0^{u_1} p_{jj}(t-u) d_u P\{\lambda \leqq s_1; \alpha \leqq u\}. \end{aligned}$$

This is equivalent to equation (2). Now let $i=j$; then $\lambda(w)$ is not the limit from the right of points of $S_i(w)$ since for a stable i the set $S_i(w)$ is the union of disjoint intervals (Theorem 1 of [2]). Hence the first term in (3) can be evaluated as follows:

$$\lim_{h\to 0} \sum_{2\leqq n\leqq [u_1/h]} \sum_{1\leqq m\leqq \min(n-1,[s_1/h])} P\{(m-1)h \leqq \lambda < mh;$$
$$x(rh) \neq i,\ m \leqq r < n,\ x(nh) = i \mid x(0) = i\} p_{ii}(t - nh)$$
$$= \lim_{h\to 0} \sum_{2\leqq n\leqq [u_1/h]} P\{\lambda \leqq \min [(n-1)h, s_1],$$
$$(n-1)h < \alpha \leqq nh \mid x(0) = i\} p_{ii}(t - nh)$$

which reduces to the same integral as before. In the above evaluations we have used the approximation of Riemann-Stieltjes sums to the Lebesgue-Stieltjes integral of a continuous function, the fact that the w-set $\{\lambda < s\}$ belongs to the Borel field generated by $\{x(t), t < s\}$, and the fact that both λ and α have continuous distributions.

By the rules of superposition of conditional probabilities (see [6, p. 36]), we have

$$P\{x(t) = j \mid \lambda\} = \int_{0-}^{t+0} P\{x(t) = j \mid \lambda, \alpha\} \big|_{\alpha=u} d_u P\{\alpha \leqq u \mid \lambda\}$$

almost everywhere on the set $\{\lambda \leqq t\}$. Evaluated at $\lambda = s \leqq t$, this becomes, by (1) and (2),

$$P\{x(t) = j \mid \lambda = s\} = \int_{s-0}^{t+0} p_{jj}(t-u) dG_{ij}(u)$$
$$= \int_{0-}^{t-s+0} p_{jj}(t - s - v) dG_{ij}(v).$$

We have thus proved that there exists a non-negative, Borel measurable (in fact, continuous from the right) function

$$r_{ij}(t) = \int_{0-}^{t+} p_{jj}(t-v) dG_{ij}(v) \tag{4}$$

such that

$$P\{x(t) = j \mid \lambda = s\} = r_{ij}(t-s) \tag{5}$$

for each $t > 0$ and all s in $(0, t)$ except a set (depending on t) of measure zero. We remark that the point of the proof above is that the conditional probability in (5) is essentially a function of $t - s$.

From (5) it follows from the definition of conditional probability that

$$p_{ij}(t) = \int_0^t q_i e^{-q_i(t-s)} r_{ij}(s) ds + \delta_{ij} e^{-q_i t}. \tag{6}$$

In the following an unspecified summation shall range over all the positive integers. Summing (6) over j, we have

$$1 = \int_0^t q_i e^{-q_i(t-s)} \sum_j r_{ij}(s) ds + e^{-q_i t}$$

or

$$\int_0^t q_i e^{-q_i(t-s)} ds = \int_0^t q_i e^{-q_i(t-s)} \sum_j r_{ij}(s) ds.$$

Hence we have

$$\sum_j r_{ij}(s) = 1 \tag{7}$$

for almost all $s>0$. Next, if $t'>0$ we have by (6),

$$p_{ij}(t+t') = \sum_k p_{ik}(t)p_{kj}(t') = \int_0^t q_i e^{-q_i(t-s)} \sum_k r_{ik}(s)p_{kj}(t')ds + e^{-q_i t}p_{ij}(t').$$

Using (6) in the last term above we have, after a simple transformation,

$$\begin{aligned} p_{ij}(t+t') &= \int_0^t q_i e^{-q_i(t-s)} \sum_k r_{ik}(s)p_{kj}(t')ds \\ &\quad + \int_{-t'}^0 q_i e^{-q_i(t-s)} r_{ij}(s+t')ds + \delta_{ij}e^{-q_i(t+t')}. \end{aligned} \tag{8}$$

On the other hand, we have directly from (6),

$$p_{ij}(t+t') = \int_{-t'}^t q_i e^{-q_i(t-s)} r_{ij}(s+t')ds + \delta_{ij}e^{-q_i(t+t')}. \tag{9}$$

Equating (8) and (9) we obtain

$$\int_0^t q_i e^{-q_i(t-s)} r_{ij}(s+t')ds = \int_0^t q_i e^{-q_i(t-s)} \sum_k r_{ik}(s)p_{kj}(t')ds.$$

Hence for each t' and t, we have

$$r_{ij}(s+t') = \sum_k r_{ik}(s)p_{kj}(t') \tag{10}$$

for almost all s in $(0, t)$. Therefore for each $t'>0$, (10) holds for almost all $s>0$.

The functions on both sides of (10) are measurable functions of the pair (s, t'). Hence the set of (s, t') for which (10) is true is measurable. By Fubini's theorem, there is a set N of measure zero with the following property: if $0<s\notin N$, then there exists a set N_s of measure zero such that if $t>s$ and $t\notin N_s$, then

$$r_{ij}(t) = \sum_k r_{ik}(s)p_{kj}(t-s). \tag{11}$$

For every $a>0$, there exists then an $s\in N$, $0<s<a$ and such that (7) holds.

Now the functions $p_{kj}(\cdot)$, $k=1, 2, \cdots$, are all uniformly continuous in $[a, \infty)$. Hence by (7) and (11) the function $r_{ij}(\cdot)$ is uniformly continuous on $[a, \infty)-N_s$. This being true for every $a>0$ it follows that $r_{ij}(\cdot)$ coincides with a continuous function almost everywhere on $(0, \infty)$. But, by (4), $r_{ij}(\cdot)$ is continuous from the right there, therefore $r_{ij}(\cdot)$ is continuous on $(0, \infty)$. Moreover $r_{ij}(0+)$ exists and equals $G_{ij}(0)$ by (4). Let us define $r_{ij}(0) = r_{ij}(0+)$.

We take this continuous function $r_{ij}(\cdot)$ to be the desired version of the conditional probability in (5); in other words, *henceforth we set unequivocally*

(5 bis) $$P\{x(t) = j \mid \lambda = s\} = r_{ij}(t - s)$$

for all $0<s<t$.

By Fubini's theorem, the equations (7) and (10) hold for each positive $s\notin M_j$ and positive $t\notin M_j(s)$, where M_j and $M_j(s)$ have measure zero for each j and s:

(7 bis) $$\sum_j r_{ij}(s) = 1;$$

(10 bis) $$r_{ij}(s + t) = \sum_k r_{ik}(s)p_{kj}(t).$$

For $s\notin M_j$, let $t_n\notin M_j(s)$, $t_n\to t$. We have, by uniform convergence,

$$r_{ij}(s + t) = \lim_{n\to\infty} r_{ij}(s + t_n) = \sum_k r_{ik}(s) \lim_{n\to\infty} p_{kj}(t_n) = \sum_k r_{ik}(s)p_{kj}(t).$$

Hence each $M_j(s)$ is empty. Let $M=\bigcup_j M_j$ and s be arbitrary. Choose any positive $s'<s$, $s'\notin M$. Then by what was just proved, we have for every k,

(12) $$r_{ik}(s) = \sum_l r_{il}(s')p_{lk}(s - s').$$

Thus we have, replacing s by $s+t$ in (12),

$$\begin{aligned} r_{ij}(s + t) &= \sum_l r_{il}(s')p_{lj}(s - s' + t) \\ &= \sum_k \sum_l r_{il}(s')p_{lk}(s - s')p_{kj}(t) = \sum_k r_{ik}(s)p_{kj}(t). \end{aligned}$$

Hence M is empty and *we have now proved that* (10 bis) *is true for all positive s and t.* Summing (10 bis) over j, we obtain

$$\sum_j r_{ij}(s + t) = \sum_k r_{ik}(s) \sum_j p_{kj}(t) = \sum_k r_{ik}(s).$$

It follows that if (7 bis) is true for a value of s then it is true for all greater values of s; *hence* (7 bis) *is true for all positive s.*

3. **Further relations and properties.** We now return to (6). The integral in the right member of (6), being that of a continuous function, has a continuous derivative for all $t>0$, hence so does the left member and we obtain

$$(13)\qquad p'_{ij}(t) = q_i[r_{ij}(t) - p_{ij}(t)], \qquad t > 0.$$

It follows from (7 bis), now valid for all $s>0$, that:

$$(14)\qquad \sum_j p'_{ij}(t) = 0, \qquad t > 0.$$

Substituting (13) into (10 bis), valid for all $s>0$, $t>0$, we obtain

$$(15)\qquad p'_{ij}(t+s) = \sum_k p'_{ik}(t)p_{kj}(s).$$

The existence and continuity of $p'_{ij}(t)$, and the results (14) and (15), all proved under the sole assumption that $q_i < \infty$, were first established by D. G. Austin in a recent paper [1]. His proof is purely analytical and elementary, and is shorter than ours if we take into account all the preliminary measure-theoretic considerations which lead up to (6). However, once this formula is written down, many will probably accept it on faith. In fact, the formulas (5 bis), (7 bis), and (10 bis) all belong to the category of the "intuitively obvious." A real advantage of our approach is that it ties together probabilistically meaningful quantities in various ways, of which we now proceed to adduce a few more.

From (4) and the continuity of $r_{ij}(\cdot)$ it follows that $G_{ij}(t)$ is continuous for $t>0$. Recalling that $F_{ij}(t)$, $G_{ij}(t)$, and $1-e^{-q_it}$ $(t\geqq 0)$ are respectively the distribution functions of α, $\alpha-\lambda$, and λ, and that the last two random variables are independent, we have

$$(16)\qquad F_{ij}(t) = \int_{0-}^{t}[1 - e^{-q_i(t-s)}]dG_{ij}(s) = \int_0^t q_ie^{-q_i(t-s)}G_{ij}(s)ds.$$

Since the last member of (16) has a continuous derivative for $t>0$, so does the first member and we obtain

$$(17)\qquad F'_{ij}(t) = q_i[G_{ij}(t) - F_{ij}(t)].$$

This sharpens a result due to Lévy [8, Lemma II.8.1] that $F_{ij}(\cdot)$ is absolutely continuous. Furthermore, it follows from (4) that at a point t where $G'_{ij}(t)$ exists, $r'_{ij}(t)$ also exists and we have

$$r'_{ij}(t) = G'_{ij}(t) + \int_{0-}^{t} p'_{jj}(t-s)dG_{ij}(s).$$

Consequently by (13) the second derivative $p''_{ij}(t)$ exists wherever $G'_{ij}(t)$ exists, that is almost everywhere since $G_{ij}(\cdot)$ is monotone. To complete the picture we add the following relation, which is implicit in (3):

$$(18)\qquad p_{ij}(t) = \int_0^t p_{jj}(t-s)dF_{ij}(s) + \delta_{ij}e^{-q_it}.$$

Upon differentiation this yields

$$(19)\qquad p'_{ij}(t) = F'_{ij}(t) + \int_0^t p'_{jj}(t-s)F'_{ij}(s)ds - \delta_{ij}q_i e^{-q_i t}.$$

Following Doob [3, Theorem 9] we write $q_{ij} = p'_{ij}(0) = \lim_{t\downarrow 0}(p_{ij}(t)/t)$, $i\neq j$; the limit exists and is finite for every i by a result of Kolmogorov [7]. We have from (13), $p'_{ij}(0+) = q_i[r_{ij}(0) - \delta_{ij}]$. The continuity of $p_{ij}(\cdot)$ in $[0, \infty)$, that of $p'_{ij}(\cdot)$ in $(0, \infty)$, and the fact that $p'_{ij}(0+)$ exists as a finite number imply by the mean value theorem that $q_{ij} = p'_{ij}(0+)$; similarly $q_i = -p'_{ii}(0+)$. Thus we have

$$(20)\qquad r_{ij}(0) = G_{ij}(0) = \frac{(1-\delta_{ij})q_{ij}}{q_i}.$$

This extends a result due to Doob [3, Theorem 10] from a stable to an arbitrary j. Since obviously $F_{ij}(0) = 0$ we have by (17) that $F'_{ij}(0+) = (1-\delta_{ij})q_{ij}$.

Turning to the limits as $t\to\infty$, we have by (16)

$$(21)\qquad F_{ij}(+\infty) = G_{ij}(+\infty).$$

Lévy [7, Theorem II.8.2] has proved that $\lim_{t\to\infty} p_{ij}(t) = \pi_{ij}$ exists for every i and j; hence it follows from (4) that $\lim_{t\to\infty} r_{ij}(t)$ exists and from (6) that $\lim_{t\to\infty} r_{ij}(t) = \pi_{ij}$. Consequently we have by (13)

$$(22)\qquad \lim_{t\to\infty} p'_{ij}(t) = 0.$$

Letting $t\to\infty$ in (10 bis) we obtain on account of (7 bis)

$$\pi_{ij} = \sum_k r_{ik}(s)\pi_{kj}, \qquad s > 0.$$

The last equation is true if $r_{ik}(\cdot)$ is replaced by $p_{ik}(\cdot)$, hence it is also true if $r_{ik}(\cdot)$ is replaced by $p'_{ik}(\cdot)$, on account of (13).

We summarize some of the results as follows.

Under the sole assumption that $0 < q_i < \infty$:

THEOREM 1. *One version of the conditional probability* $P\{x(t) = j \mid \lambda = s\}$ *is a continuous function* $r_{ij}(t-s)$ *for* $0 < s < t < \infty$ *which satisfies the equations* (6), (7 *bis*), *and* (10 *bis*) *for all* $s > 0$ *and* $t > 0$.

THEOREM 2. *The distribution function* $G_{ij}(\cdot)$ *of* $\alpha_{ij} - \lambda_i$ *is a continuous function in* $[0, \infty)$ *satisfying* (20) *and* (21). *The distribution function* $F_{ij}(\cdot)$ *of* α_{ij} *has a continuous derivative in* $[0, \infty)$ *given by* (17).

THEOREM 3. *The function* $p_{ij}(\cdot)$ *has a continuous derivative in* $[0, \infty)$ *given by* (13) *satisfying* (14), (15), *and* (22). *The second derivative exists as a finite number almost everywhere.*

THEOREM 4. *The functions $p_{ij}(\cdot)$, $r_{ij}(\cdot)$, $F_{ij}(\cdot)$, and $G_{ij}(\cdot)$ are related by* (4), (6), (16), *and* (18).

4. **The post-exit process.** We define a new process $\{y(t),\ t>0\}$, as follows:

$$y(t) = y(t, w) = x(\lambda_i(w) + t, w), \qquad t > 0. \tag{23}$$

We shall see later that this indeed yields a process which may be taken as separable relative to the closed sets and measurable. This process will be called a *post-exit process, derived from the Markov chain $x(t)$ and the stable state i.* The following theorem shows that as far as its transition is concerned, it is exactly like the original $x(t)$ process. This transparent result proves to be surprisingly elusive of proof; no doubt this is due to the fact the sample function behavior to the immediate right of $\lambda_i(w)$ is not simple, see the Complement to the theorem below.

THEOREM 5. *Let $\{y(t);\ t>0\}$ be the post-exit process defined by* (23). *Let $0<t_1<t_2< \cdots <t_l$, and $j_1, j_2, \cdots, j_l$ be arbitrary states. Then we have*

$$P\{y(t_\nu) = j_\nu, 1 \leqq \nu \leqq l\} = r_{ij_1}(t_1) \prod_{\nu=1}^{l-1} p_{j_\nu j_{\nu+1}}(t_{\nu+1} - t_\nu).$$

Proof. We need some general information regarding the sample functions of the $x(t)$ process. Detailed theorems of this kind can be found in [2], but the relevant facts here can be briefly described as follows. Let $S_j(w) = \{t: x(t, w) = j\}$. First, it is true with probability one that every point of $S_j(w)$ is a limit point of $S_j(w)$(4). Next, by a fundamental theorem of Doob [3, Theorem 12], as $t \to \tau(w)$(5) from one side there is at most one finite limiting value of $x(t, w)$, with probability one. Moreover, it is possible to make the separable process $x(t)$ so that $x(\tau(w), w) = +\infty$ if and only if $\lim_{t\to\tau(w)} x(t, w) = +\infty$. Consequently if $\tau(w)$ is a limit point of $S_i(w)$ from both sides, then $x(\tau(w), w) = i$. We have thus proved the following inclusions with probability one

$$\bigcap_m \bigcap_{\nu=1}^{l} \{S_{j_\nu} \cap (\lambda + t_\nu - m^{-1}, \lambda + t_\nu) \neq 0, S_{j_\nu} \cap (\lambda + t_\nu, \lambda + t_\nu + m^{-1}) \neq 0\}$$

$$\subseteq \bigcap_{\nu=1}^{l} \{x(\lambda + t_\nu) = j_\nu\}$$

$$\subseteq \bigcap_m \bigcap_{\nu=1}^{l} \{S_{j_\nu} \cap (\lambda + t_\nu - m^{-1}, \lambda + t_\nu + m^{-1}) \neq 0\}.$$

(4) See footnote (3).

(5) For lack of a better notation, we write $\tau(w)$ to denote a generic abscissa in the "graph" $(\tau, x(\tau, w))$ corresponding to the generic w. Thus it is neither fixed irrespective of w nor properly speaking a function of w.

We prove that the last-written w-set above is measurable as follows. Let $f(t, w)=1$ if $t\in S_j(w)$ and $=0$ otherwise; let $g(t, w)=1$ if $\lambda(w)+t_\nu-m^{-1}<t<\lambda(w)+t_\nu+m^{-1}$ and $=0$ otherwise. Then both f and g are measurable functions of the pair (t, w). Thus the function fg is measurable in the pair and by Fubini's theorem for almost all s, $f(s, w)g(s, w)$ is a measurable function of w. Thus there exists a denumerable and everywhere dense t-set $\{s_n\}$ such that $\sum_n f(s_n, w)g(s_n, w)$ is a measurable function of w. This set $\{s_n\}$ may be taken to be the set satisfying the conditions of the separability definition of the $x(t)$ process, since $\lim_{t\to 0} p_{ij}(t)=\delta_{ij}$ [6, Theorem II.2.2]. We have

$$\{S_{j_\nu}\cap(\lambda+t_\nu-m^{-1}, \lambda+t_\nu+m^{-1})\neq 0\}=\left\{\sum_n f(s_n, w)g(s_n, w)>0\right\}$$

which is therefore a measurable w-set, as was to be proved.

Let $0<s<t_1$. We have, setting $h=2^{-p}$, $p\to\infty$,

$$\begin{aligned}
P&\left\{\bigcap_m\bigcap_{\nu=1}^{l}[S_{j_\nu}\cap(\lambda+t_\nu-m^{-1}, \lambda+t_\nu+m^{-1})\neq 0]\right\}\\
&\leqq \lim_{m\to\infty}\lim_{h\to 0}\sum_{n=0}^{\infty}\sum_j P\{nh\leqq\lambda<(n+1)h, x(nh+h+s)=j\}\\
&\quad\cdot P\{S_{j_\nu}\cap(t_\nu+nh-m^{-1}, t_\nu+nh+h+m^{-1})\neq 0, 1\leqq\nu\leqq l \mid x(nh+h+s)=j\}\\
&= \lim_{m\to\infty}\lim_{h\to 0}\sum_{n=0}^{\infty}\sum_j P\{nh\leqq\lambda<(n+1)h, x(nh+h+s)=j\}\\
&\quad\cdot P\{S_{j_\nu}\cap(t_\nu-s-h-m^{-1}, t_\nu-s+m^{-1})\neq 0, 1\leqq\nu\leqq l \mid x(0)=j\}\\
&\leqq \lim_{h\to 0}\sum_{n=0}^{\infty}\sum_j P\{nh\leqq\lambda<(n+1)h, x(nh+h+s)=j\}\\
&\quad\cdot\lim_{m\to\infty} P\{S_{j_\nu}\cap(t_\nu-s-2m^{-1}, t_\nu-s+m^{-1})\neq 0, 1\leqq\nu\leqq l \mid x(0)=j\}\cdot
\end{aligned}$$

The last step is justified as follows:

$$\begin{aligned}
\sum_{n=0}^{\infty}\sum_j &P\{nh\leqq\lambda<(n+1)h, x(nh+h+s)=j\}\\
&= \sum_{n=0}^{\infty}\sum_j\int_{nh}^{(n+1)h} q_i e^{-q_i u} r_{ij}(nh+h+s-u)du\\
&= \sum_j\int_0^{\infty} q_i e^{-q_i u} r_{ij}([uh^{-1}]h+h+s-u)du.
\end{aligned}$$

Since $r_{ij}(\cdot)$ is non-negative and continuous, the convergence in (7bis) is uniform with respect to s in every $[a, b]$ where $0<a<b<\infty$. Hence the last series above converges uniformly with respect to h in $(0, 1)$. This is sufficient for the interchange of limits in the last step.

Similarly, we have

$$P\left\{\bigcap_m \bigcap_{\nu=1}^{l} [S_{j_\nu} \cap (\lambda + t_\nu - m^{-1}, \lambda + t_\nu) \neq 0; S_{j_\nu} \cap (\lambda + t_\nu, \lambda + t_\nu + m^{-1}) \neq 0]\right\}$$
$$\geq \lim_{h\to 0} \sum_{n=0}^{\infty} \sum_j P\{nh \leq \lambda < (n+1)h,\ x(nh + h + s) = j\}$$
$$\cdot \lim_{m\to\infty} P\{S_{j_\nu} \cap (t_\nu - s - m^{-1}, t_\nu - s - 2^{-1}m^{-1}) \neq 0,$$
$$S_{j_\nu} \cap (t_\nu - s, t_\nu - s + 2^{-1}m^{-1}) \neq 0,\ 1 \leq \nu \leq l \mid x(0) = j\}\cdot$$

Now for fixed t_ν, it follows from a theorem of Doob [4, p. 457] (see Theorem 2 of [2]) that the two limits in m at the end of the last two sets of inequalities have the same value given by

$$p_{jj_1}(t_1 - s) \prod_{\nu=1}^{l-1} p_{j_\nu j_{\nu+1}}(t_{\nu+1} - t_\nu) = p_{jj_1}(t_1 - s)Q$$

where Q stands for the product in ν. Hence we have, since our basic probability is tacitly assumed to be a complete measure:

$$P\{y(t_\nu) = j_\nu\}$$
$$= \lim_{h\to 0} \sum_{n=0}^{\infty} \sum_j P\{nh \leq \lambda < (n+1)h,\ x(nh + h + s) = j\} p_{jj_1}(t_1 - s)Q.$$

Using (5) and (10 bis), we have

$$P\{y(t_\nu) = j_\nu\} = Q \lim_{h\to 0} \sum_{n=0}^{\infty} \sum_j \int_{nh}^{(n+1)h} q_i e^{-q_i u} r_{ij}(nh + h + s - u) p_{jj_1}(t_1 - s)du$$
$$= Q \lim_{h\to 0} \sum_j \int_0^\infty q_i e^{-q_i u} r_{ij}([uh^{-1}]h + h + s - u) p_{jj_1}(t_1 - s)du$$
$$= Q \lim_{h\to 0} \int_0^\infty q_i e^{-q_i u} r_{ij_1}([uh^{-1}]h + h + t_1 - u)du$$
$$= Q \int_0^\infty q_i e^{-q_i u} r_{ij_1}(t_1)du = r_{ij_1}(t_1)Q$$

since $0 \leq r_{ij}(\cdot) \leq 1$ and $r_{ij}(\cdot)$ is continuous. This proves the theorem.

COROLLARY 1. *There is a standard modification of the $y(t)$ process which is separable relative to the class of closed sets and measurable.*

Proof. By Theorem II.2.6 of [6], it is sufficient to prove that for every $t>0$ and $\epsilon>0$

$$\lim_{s\to t} P\{\,|\,y(t) - y(s)\,| < \epsilon\} = 1.$$

Now we have by the theorem, if $0<\epsilon<1$,

$$\begin{aligned}\lim_{s\uparrow t} P\{\,|\,y(t)-y(s)\,|\,<\epsilon\} &= \lim_{s\uparrow t}\sum_j P\{y(s)=y(t)=j\}\\ &= \lim_{s\uparrow t}\sum_j r_{ij}(s)p_{jj}(t-s) \geqq \sum_j r_{ij}(t) = 1\end{aligned}$$

by (7 bis); similarly for $s\downarrow t$.

COROLLARY 2. *For each* $t>0$, $P\{y(t)=+\infty\}=0$. *Thus the* $y(t)$ *process as well as the* $x(t)$ *process has no "fixed infinities."*

COMPLEMENT TO THEOREM 5. *For every stable* j, *we have*

$$P\left\{\lim_{t\downarrow 0} y(t) = j\right\} = \frac{(1-\delta_{ij})q_{ij}}{q_i};$$

for every instantaneous j, *we have*

$$P\left\{\liminf_{t\downarrow 0} y(t) = j < +\infty = \limsup_{t\downarrow 0} y(t)\right\} = \frac{q_{ij}}{q_i};$$

$$P\left\{\lim_{t\downarrow 0} y(t) = +\infty\right\} = 1 - \frac{\sum_{j\neq i} q_{ij}}{q_i}.$$

The Complement follows from (20) and Theorem 4 of [2]. Examples of all three kinds of behavior can be found in Kolmogorov [7] and Lévy [8, pp. 365–367].

5. **A counterpart.** The exact counterpart to a post-exit process would be a *pre-entrance process*. We shall not introduce this formally but choose the following shorter course.

Let j be a stable state. Whatever the initial distribution of $x(0)$ may be, the t-set $S_j(w)=\{t:x(t,w)=j\}$ consists of a sequence of disjoint intervals, called the j-intervals, of which there is a finite number in every finite t-interval, with probability one. The total number of j-intervals may be finite or infinite. The beginning of the nth j-interval is a random variable (in the broad sense) denoted by $\tau_j^n=\tau_j^n(w)$ and called the nth entrance time into the state j; the length of the nth j-interval is a random variable (in the broad sense) denoted by $\lambda_j^n=\lambda_j^n(w)$ and called the nth sojourn time in the state j. If $P\{x(0)=i\}=1$ then we have $P\{\tau_j^1\leqq t\}=F_{ij}(t)$; $P\{\tau_j^{n+1}-\tau_j^n\leqq t\}=F_{jj}(t)$(6), $n=1,2,\cdots$; and $P\{\lambda_j^n\leqq t\}=1-e^{-q_jt}$, $t\geqq 0$. The random variables τ_j^n and λ_j^n are independent; and the sequence of random variables $\{\tau_j^{n+1}-\tau_j^n\}$, $n=0, 1, 2, \cdots$, where $\tau_j^0=0$, are independent. All the preceding statements can be found in Lévy's work [8]; proofs can be found in [2].

We have, clearly, if $t>0$,

(6) Note that if $q_j=0$ then $F_{jj}(\ )\equiv 0$.

$$P\{x(t) = j\} = \sum_{n=1}^{\infty} P\{\tau_j^n \leqq t < \tau_j^{n+1}, x(t) = j\}$$
$$= \sum_{n=1}^{\infty} P\{\tau_j^n \leqq t, \lambda_j^n > t - \tau_j^n\}.$$

Hence we have, by the independence of τ_j^n and λ_j^n,

$$P\{x(t) = j\} = \sum_{n=1}^{\infty} \int_0^{t+0} e^{-q_j(t-s)} dP\{\tau_j^n \leqq s\}. \tag{24}$$

This formula has been given by Lévy [8, (II.8.1)]. Let $P\{x(0)=i\}=1$, then

$$\sum_{n=1}^{\infty} P\{\tau_j^n \leqq s\} = U_{ij}(s),$$

where

$$U_{ij}(s) = \delta_{ij}\epsilon(s) + \sum_{n=0}^{\infty} F_{ij}(s) * F_{jj}^{n*}(s)$$

and F^{n*} denotes the n-fold convolution of F with itself, $F_{jj}^{0*}(s)$ being the unitary distribution function $\epsilon(s)$. The function $U_{ij}(s)$, which is finite for finite s and nondecreasing, is well-known in renewal theory and may be interpreted as the expected number of entrances into j in time s (see [5]).

So far the initial state i is arbitrary. We now assume that i is a stable state. Then by Theorem 2 both $F_{ij}(\cdot)$ and $F_{jj}(\cdot)$ have continuous derivatives. It follows easily that the derivative

$$u_{ij}(s) = [U_{ij}(s) - \delta_{ij}\epsilon(s)]' = \int_0^s F'_{ij}(s-v)d\left[\sum_{n=0}^{\infty} F_{jj}^{n*}(v)\right] \tag{25}$$

exists and is non-negative and continuous for all $s>0$. We can now rewrite (24) as

$$p_{ij}(t) = \int_0^t e^{-q_j(t-s)} u_{ij}(s)ds + \delta_{ij}e^{-q_jt}. \tag{26}$$

Differentiating, we obtain

$$p'_{ij}(t) = u_{ij}(t) - q_j p_{ij}(t) \tag{27}$$

which is the counterpart to (13). Note however that (27) *is proved under the double condition*: $q_i < \infty$ and $q_j < \infty$.

Henceforth we shall assume the more stringent condition that $q_k < \infty$ *for every* k. Writing k for i in (26), multiplying by $p_{ik}(t')$ and summing over k, we obtain, for every $t'>0$:

$$
\begin{aligned}
\sum_k p_{ik}(t')p_{kj}(t) &= \int_0^t e^{-q_j(t-s)} \sum_k p_{ik}(t')u_{kj}(s)ds + p_{ij}(t')e^{-q_jt} \\
&= \int_0^t e^{-q_j(t-s)} \sum_k p_{ik}(t')u_{kj}(s)ds \\
&\quad + \int_{-t'}^0 e^{-q_j(t-s)}u_{ij}(t'+s)ds + \delta_{ij}e^{-q_j(t'+t)}.
\end{aligned} \tag{28}
$$

On the other hand, we have

$$
p_{ij}(t'+t) = \int_{-t'}^t e^{-q_j(t-s)}u_{ij}(t'+s)ds + \delta_{ij}e^{-q_j(t'+t)}. \tag{29}
$$

Equating (28) and (29), we have

$$
\sum_k p_{ik}(t)u_{kj}(s) = u_{ij}(t+s) \tag{30}
$$

for each $t>0$ and almost all $s>0$. Since both members of (30) are measurable functions of the pair (t, s), it follows by Fubini's theorem that there are sets N and N_s of measure zero such that if $0<s\notin N$ and $0<t\notin N_s$ then (30) holds for all i and j. Furthermore since the functions $u_{kj}(\cdot)$ are non-negative and continuous, it follows from (30) that

$$
\sum_k p_{ik}(t)u_{kj}(s) \leqq u_{ij}(t+s) \tag{31}
$$

for all positive t and s. Suppose that for some t and s strict inequality holds in (31), then we have since $p_{ii}(t')>0$ for every $t'>0$,

$$
\begin{aligned}
\sum_k p_{ik}(t'+t)u_{kj}(s) &= \sum_l p_{il}(t') \sum_k p_{lk}(t)u_{kj}(s) \\
&< \sum_l p_{il}(t')u_{lj}(t+s) \leqq u_{ij}(t'+t+s).
\end{aligned}
$$

Thus if (30) does not hold for certain s and t, then it does not hold for the same s and all greater t. Thus the sets N_s for $s\notin N$ must be all empty. Let s and t be positive but arbitrary. Choose $s'<s$ such that $s-s'\notin N$. Then we have just proved that

$$
u_{kj}(t+s-s') = \sum_l p_{kl}(t')u_{lj}(s-s') \tag{32}
$$

for all k and j. Hence we have, using (32) with s' resp. $t+s'$ for t,

$$
\begin{aligned}
\sum_k p_{ik}(t)u_{kj}(s) &= \sum_l \sum_k p_{ik}(t)p_{kl}(s')u_{lj}(s-s') \\
&= \sum_l p_{il}(t+s')u_{lj}(s-s') = u_{ij}(t+s).
\end{aligned}
$$

Therefore (30) holds in fact for all positive t and s. The relation (30) may be interpreted as a transition property of the "entrance frequency" $u_{ij}(s)$.

Substituting (27) into (30), we obtain(7)

$$p'_{ij}(t+s) = \sum_k p_{ik}(t)p'_{kj}(s). \tag{33}$$

This is the exact counterpart to (15), proved however under the more stringent condition that all $q_k < \infty$.

We summarize the results as follows.

THEOREM 6. *If $q_i < \infty$ and $q_j < \infty$ then the function $u_{ij}(s)$ given by* (25) *is continuous for all $s > 0$ and satisfies* (26) *and* (27). *If $q_k < \infty$ for every k then the relations* (30) *and* (33) *hold.*

6. The generalized Kolmogorov equations. The pair of equations (15) and (33):

$$p'_{ij}(t+s) = \sum_k p'_{ik}(s)p_{kj}(t), \quad p'_{ij}(t+s) = \sum_k p_{ik}(t)p'_{kj}(s)$$

may be called the *generalized Kolmogorov equations*. The first equation (obtained first by Austin) is proved under the hypothesis that $q_i < \infty$, the second under the hypothesis that $q_k < \infty$ for every k; they are both valid for all positive t and s. Since $p'_{ij}(\cdot)$ is a continuous function, we have as $s \downarrow 0$:

$$p'_{ij}(t) = \lim_{s \downarrow 0} \sum_k p'_{ik}(s)p_{kj}(t), \qquad t > 0,$$

$$p'_{ij}(t) = \lim_{s \downarrow 0} \sum_k p_{ik}(t)p'_{kj}(s), \qquad t > 0.$$

The historical approach is equivalent to attempts to take the limits above underneath the summation sign, and it is found necessary to impose further conditions to make this legitimate (see [4]). From our point of view the equations (15) and (33) are consequences of (10 bis) and (30), respectively, which are probabilistically more meaningful. They have the further analytical advantage of involving only non-negative quantities. To illustrate this we prove the following theorem which sharpens a result of Doob [4, p. 461].

THEOREM 7. *Suppose that $q_i = -p'_{ii}(0) < \infty$. If the equation*

$$p'_{ij}(t) = \sum_k p'_{ik}(0)p_{kj}(t) \tag{34}$$

is true for some $t > 0$ and every j, then

$$\sum_k p'_{ik}(0) = 0; \tag{35}$$

(7) Although the equation (33) is given in Austin's paper [1], according to an oral communication his proof is incomplete.

conversely if (35) *is true, then* (34) *is true for every* $t>0$ *and* j. *Furthermore, either assertion is equivalent to the following assertion regarding the post-exit process* (23):

$$P\left\{\lim_{t\downarrow 0} y(t) = +\infty\right\} = 0.$$

Proof. (34) and (35) are equivalent, by means of (13), to the following limit forms of (10 bis) and (7 bis):

$$r_{ij}(t) = \sum_k r_{ik}(0)p_{kj}(t), \tag{36}$$

$$\sum_k r_{ik}(0) = 1. \tag{37}$$

If (36) is true for some $t>0$ and every j, then summing over j and using (7 bis) we have $1=\sum_j r_{ij}(t)=\sum_k r_{ik}(0)\sum_j p_{kj}(t)=\sum_k r_{ik}(0)$. Conversely, suppose (37) is true. By letting $s\to 0$ in (10 bis) we have for every $t>0$

$$r_{ij}(t) \geqq \sum_k r_{ik}(0)p_{kj}(t). \tag{38}$$

Hence we have as before

$$1 = \sum_j r_{ij}(t) \geqq \sum_k r_{ik}(0) = 1.$$

Thus we must have equality in (38) for every j.

The last statement follows from the Complement to Theorem 5, since (35) is the same as $\sum_{j\neq i} q_{ij}=q_i$.

References

1. D. G. Austin, *On the existence of derivatives of Markoff transition probability functions*, Proc. Nat. Acad. Sci. U.S.A. vol. 41 (1955) pp. 224–226.
2. K. L. Chung, *Foundations of the theory of continuous parameter Markov chains*, to appear in the Third Berkeley Symposium.
3. J. L. Doob, *Topics in the theory of Markoff chains*, Trans. Amer. Math. Soc. vol. 52 (1942) pp. 37–64.
4. ———, *Markoff chains—denumerable case*, Trans. Amer. Math. Soc. vol. 58 (1945) pp. 455–473.
5. ———, *Renewal theory from the point of view of the theory of probability*, Trans. Amer. Math. Soc. vol. 63 (1948) pp. 422–438.
6. ———, *Stochastic processes*, New York, 1953.
7. A. N. Kolmogorov, *On the differentiability of the transition probabilities in stationary Markov processes with a denumerable number of states*, Ucenye Zapiski Moskovskogo Gosudarstvennogo Universiteta. Matematika vol. 4 (1951) pp. 53–59 (Russian).
8. P. Lévy, *Systèmes Markoviens et stationnaires. Cas dénombrable*, Ann. École Norm. vol. 68 (1951) pp. 327–381.

Syracuse University,
Syracuse, N. Y.

Reprinted with kind permission.

ANNALS OF MATHEMATICS
Vol. 68, No. 1, July, 1958
Printed in Japan

ON A BASIC PROPERTY OF MARKOV CHAINS[1]

BY K. L. CHUNG

(Received June 27, 1957)
(Revised November 12, 1957)

1. The main question discussed in this paper has been in the air for some time. Roughly speaking it may be stated as follows: Given a continuous parameter[2] Markov process $\{x_t, t \geqq 0\}$ with stationary transition probabilities, let us start it off at a randon time α chosen according to the way it has been unravelling, as it were, but without prevision of the future. Does the shifted process $\{x_{\alpha+t}, t \geqq 0\}$ remain Markovian with the same transition probabilities, and, knowing x_α, is the past $\{x_t, t \leqq \alpha\}$ independent of the future $\{x_t, t \geqq \alpha\}$? Observe that if the random time is in fact a constant, then the required property simply reduces to the Markovian and stationarity hypothesis. Hence the new property, in various versions[3], has been called the *strong Markovian property*. Familiar examples of the *optional starting time* are the first entry into, or exit from a given set.[4] Such special cases of optional starting have been found of great importance in the study of Markov processes, particularly in the Brownian motion process and the so-called diffusion process. In the general case, however, the answer to the preceding quetion is negative : indeed a counter-example is obtained by modifying the transition probability function of a Brownian motion process at a single point. For the Brownian motion process itself, the answer is affirmative, but here it must be said that the formulation of the problem is the better half of its solution, owing to the very special nature of the process. Matters at the other end of the scale, in the chain case, namely where the state space is denumerable, are not so simple. If no extraneous assumptions are made, it has been shown by Feller and McKean [7] that all the states may be instantaneous so that the typical sample function may take on distinct values on a denumerable number of Cantor-like sets

[1] Research carried out at the Statistical Research Center, University of Chicago, under sponsorship of the Statistics Branch, Office of Naval Research. Reproduction in whole or in part is permitted for any purpose of the United States Government.

[2] For a discrete-parameter process the answer is affirmative and trivial.

[3] Previous results (see [0], [6], [9], [11]) for a euclidean space but based on the right continuity of sample functions, the finiteness of $x(\alpha(w), w)$ etc, are irrelevant here since these assumptions drastically simplify the problem. The definition of the strong Markov property given in [6] is the closest to the one given below in Remark 1 to Theorem 3; however it is not explained in [6] what $P_{x(\tau)}$ means when $x(\tau) = \infty$ (cf. Theorem 5 below).

[4] Related notions play an important role in Doob's theory of martingales.

which fill up almost the all time axis without any interval of constancy — a state of affairs remote from the common notions of continuity. Even if all the states are stable, the optional time may occur at an instant τ when the sample function $x(t, w)$ explodes to infinity both for $t \downarrow \tau$ and for $t \uparrow \tau$ and again no kind of facile continuity is available.[5] Yet one realizes that the question is one of approximating the random shift by means of constant shifts where a sort of continuity argument is the inevitable issue.

In this paper an affirmative answer to the proposed question, to be made precise in Theorems 3 and 5 below, will be given in the chain case, without any extraneous hypothesis. The problem as such was first put to me by Doob. In fact he gave the first simple result along this line in 1945 [3 ; Theorem 3.1], concering the entry into a stable state. After Lévy's theory of sample functions [10] appeared, it was an easy matter to extend this to an instantaneous state [1 ; Theorem 6]. The next step taken concerns the exit from a stable state [2 ; Theorem 5]. This result already contained, as I have come to realize, the broad feature of the general situation. The method used here is essentially a generalization of that used there, but no previous knowledge of that paper is required. It may be pointed out that the main difference here is that a certain function of (s, t) is no longer a continuous function of the difference $t - s$, as it was in the post-exit case, but fortunately its continuity in t for each s suffices. The whole question is on a rather delicate level and I now find my previous exposition frequently too brief. This time I have taken pains to be more expansive. Two theorems concerning the Borel measurability of a general stochastic process are given at the beginning although, as in [2], this is not indispensable. A moot point about separability is mentioned toward the end of the paper.[6] The implications of the main result, particularly as applied to the study of the so-called "boundaries," cannot yet be fully discussed, since the latter theory is far from being ready. I hope to take up this topic at a later time.

I am greatly indebted to Henry McKean, Jr. for important revisions of my manuscript.

2. Let T denote the t-interval $[0, \infty)$; T^0 the t-interval $(0, \infty)$; I the set of positive integers ; $\bar{I} = I \cup \{\infty\}$, more exactly $\bar{I}$ is the one-point compactification of I in the discrete topology. The letters i, j, k, m, n will stand for elements of I ; the letters s, t, u will stand for points of T.

[5] These illustrations should warn those unfamiliar with the chain case against drawing random conclusions.

[6] Added in proof: This has now been resolved; see Addenda.

Let $(\Omega, \mathscr{F}, \mathbf{P})$ be a probability triple where the couple $(\mathscr{F}, \mathbf{P})$ is *complete* in the sense that if $\Lambda \in \mathscr{F}$ and $\mathbf{P}(\Lambda) = 0$ then every subset of Λ belongs to $\mathscr{F}$. Let $\{\hat{x}(t, w), t \in T\}$, where w denotes a generic point of Ω, be a continuous parameter Markov chain with stationary transition probabilities defined on the triple. We assume that for each t,

$$\mathbf{P}\{\hat{x}(t, w) \in \bar{I}\} = 1 ;$$

furthermore, we assume that for each $j \in I$, there exists a $t > 0$ for which $\mathbf{P}\{\hat{x}(t, w) = j\} > 0$. The set $\bar{I}$ is then called *the state space* of the Markov chain and each element of $\bar{I}$ a *state*. The *adjoined state* ∞ obviously plays a special role. Let us denote the initial probabilities by

$$p_i = \mathbf{P}\{\hat{x}(0, w) = i\} , \qquad i \in \bar{I};$$

the transition probabilities by

$$p_{ij}(t) = \mathbf{P}\{\hat{x}(s + t, w) = j | \hat{x}(s, w) = i\} , \quad i, j \in \bar{I}, s, t \in T ;$$

and the absolute probabilities by

$$p_j(t) = \mathbf{P}\{\hat{x}(t, w) = j\} = \textstyle\sum_{i \in \bar{I}} p_i p_{ij}(t) , \qquad j \in \bar{I}, t \in T.$$

For the adjoined state ∞ we put

$$p_\infty = 0, \; p_{i\infty}(t) = p_{\infty i}(t) = \delta_{i\infty} \qquad i \in \bar{I}, t \in T ;$$

so that $p_\infty(t) = 0$ for all $t \in T$, in contrast with the other states. The other probabilities are characterized by the following conditions :

$$p_i \geqq 0, \qquad \textstyle\sum_{i \in \bar{I}} p_i = 1 ; \qquad i \in I$$

$$p_{ij}(t) \geqq 0, \qquad \textstyle\sum_{j \in \bar{I}} p_{ij}(t) = 1 ; \qquad i, j \in I, t \in T$$

$$p_{ij}(s+t) = \textstyle\sum_{k \in \bar{I}} p_{ik}(s) p_{kj}(t), \qquad i, j \in I, s, t \in T.$$

The range of summation $\bar{I}$ in the equations may be replaced by I ; as they stand the six relations above remain true if $i, j \in \bar{I}$.

So far we have not restricted our chain in any essential way (except the compactification[7] of I). The following minimum regularity assumption will now be imposed :

$$(1) \qquad \lim_{t \downarrow 0} p_{ii}(t) = 1 \qquad i \in I.$$

This assumption is equivalent to the stochastic continuity of the process at each fixed t. It is known that there is always a standard modification of the x-process that is separable relative to the closed sets (henceforth "separable" for short). The assumption (1) is, under

[7] Recent work points to the desirability of a more refined way of compactifcation, leading to various adjoined states (boundaries) instead of the single ∞. Fortunately this will not be necessary in the present paper.

our definition of $\bar{I}$, equivalent to the existence of a separable and measurable standard modification [3 ; Theorem 8]. Then any denumerable set dense in T may be taken to be the set satisfying the conditions of the separability definition (henceforth "a separability set" for short); see [5, pp. 54, 59]. We may and will suppose the $\hat{x}$-process itself to be separable and measurable. For each $w \in \Omega$, the sample function $\hat{x}(\circ, w)$ maps T into $\bar{I}$; in fact separability requires a compact range obtained by adjoining the state ∞ to I.

The following statement [1 ; Theorem 2] is then true. For each fixed t, we have

$$(2) \qquad \mathbf{P}\{\underline{\lim}_{s\to t}\hat{x}(s, w) = \hat{x}(t, w)\} = 1 .$$

In other words, almost all sample functions $x(\circ, w)$ are lower semi-continuous at each fixed t. The statement remains true even if we replace $s \to t$ in (2) by $s \downarrow t$ or $s \uparrow t$; see (2′) below.

3. We shall now give a general result applicable to any stochastic process satisfying the condition (2). Recall that the process $\{x(t,w), t \in T\}$ is said to be measurable in case x as a function of the pair (t, w) is measurable in the completed Borel field $\overline{\mathscr{B} \times \mathscr{F}}$ where $\mathscr{B}$ is the classical Borel field on T.

DEFINITION 1. The stochastic process $\{x(t, w), t \in T\}$ is called Borel measurable in case x is measurable in the product field $\mathscr{B} \times \mathscr{F}$ (without completion).

This more strict concept is often useful. It guarantees that for each w the sample function $x(\circ, w)$ is measurable $\mathscr{B}$, namely Borel measurable in the classical sense, or again a Baire function. In particular if x and y are measurable $\mathscr{B} \times \mathscr{F}$, then the substitution of y for the t in $x(t, w)$ will yield a $z(t, w) = x(y(t, w), w)$ that is measurable $\mathscr{B} \times \mathscr{F}$. Thus for each t, $z(t, \circ)$ will be measurable $\mathscr{F}$, namely a random variable. We shall use a particular case of this in the sequel.

THEOREM 1.[8] *If an arbitrary separable stochastic process $\{\hat{x}(t, w), t \in T\}$ satisfies the condition* (2) *for each fixed* t, *then there is a standard modification* $\{\tilde{x}(t, w), t \in T\}$, *defined by* $\tilde{x}(t, w) = \underline{\lim}_{s\to t}\hat{x}(s, w)$, *that is separable*

[8] As pointed out by Austin, the lower limits in the proofs are taken in a neighborhood including t. However, Theorems 1 and 2 remain true if the lower limits are taken in deleted neighborhoods; thus our lower semi-continuity is more exigeant than defined, e. g., in [8; p. 307] The proofs need only slight changes owing to separability. This remark is essential for later applications.

with the same separability set, and Borel measurable. Almost all sample functions $\tilde{x}(\circ, w)$ are lower semi-continuous in T.

REMARK. It goes without saying that the $\underline{\lim}$ in (2) may be replaced by $\overline{\lim}$ with the proper changes.

PROOF. Let $\{s_j, j \in I\}$ be a separability set. Define

$$x^{(n)}(t, w) = \inf_{\frac{2m-1}{2^{n+1}} \leq s < \frac{2m+3}{2^{n+1}}} \hat{x}(s, w) \qquad \text{if } \frac{m}{2^n} \leqq t < \frac{m+1}{2^n}.$$

Because of separability the inf need only be taken over a denumerable set, hence it is measurable $\mathscr{F}$ and consequently $x^{(n)}$ is measurable $\mathscr{B} \times \mathscr{F}$. For each (t, w), $x^{(n)}(t, w)$ is non-decreasing and approaches $\tilde{x}(t, w)$ as $n \to \infty$. It follows from (2) that for each t,

$$\mathbf{P}\{\tilde{x}(t, w) = \hat{x}(t, w)\} = 1 ,$$

namely that $\tilde{x}$ is a standard modification of $\hat{x}$. Furthermore for each s_j there is a set Z_j of probability zero such that if $w \notin Z_j$ then $\tilde{x}(s_j, w) = \hat{x}(s_j, w)$. Consequently if $w \notin Z = \bigcup_j Z_j$, this equation holds simultaneously for all $j \in I$. If G is an open interval containing t, then by its definition the value of $x^{(n)}(t, w)$ is in the closure of the values of $\hat{x}(s, w)$, $s \in G$; hence of $\hat{x}(s_j, w)$, $s_j \in G$; hence also of $\tilde{x}(s_j, w)$, $s_j \in G$, provided that $w \notin Z$. This being so for every n the same is true of the limit $\tilde{x}(t, w)$. Therefore the $\tilde{x}$-process is separable with $\{s_j\}$ as a separability set. The last assertion of Theorem 1 follows from the fact that

$$\underline{\lim}_{s \to t} \tilde{x}(s, w) = \tilde{x}(t, w) .$$

For the purposes of this paper we shall need the following more precise result. Theorem 1 itself is not used in this paper but is given separately for its general interest.

THEOREM 2. *Suppose that the separable stochastic process $\{\hat{x}(t, w), t \in T\}$ satisfies the following condition: for each fixed t we have*

$$\text{(2')} \qquad \mathbf{P}\{\underline{\lim}_{s \uparrow t} \hat{x}(s, w) = \underline{\lim}_{s \downarrow t} \hat{x}(s, w) = \hat{x}(t, w)\} = 1 .$$

(For $t = 0$ the first $\underline{\lim}$ is to be omitted.) *Then the process $\{x(t, w), t \in T\}$, where $x(t, w) = \underline{\lim}_{s \downarrow t} \hat{x}(s, w)$, is a separable (with the same separability set) and Borel measurable standard modification of the $\hat{x}$-process. For the x-process, almost all sample functions are right lower semi-continuous in T.*

PROOF. Let $\tilde{x}$ be as in Theorem 1; then it is an elementary fact that

$$x(t, w) = \underline{\lim}_{s \downarrow t} \tilde{x}(s, w) .$$

The x-process is a standard modification of the $\hat{x}$-process on account of the condition (2′); it is separable for the same reason that $\tilde{x}$ is, and has the same separability set as $\hat{x}$. It remains to prove that x is Borel measurable.[9] Let $\{s_j\}$ be a separability set; then we have

$$x(t, w) = \underline{\lim}_{s_j \downarrow t} \tilde{x}(s_j, w) = \lim_{n \to \infty} x^{(n)}(t, w)$$

where

$$x^{(n)}(t, w) = \inf_{t < s_j < t+n^{-1}} \tilde{x}(s_j, w).$$

Given a real number c let $\tilde{x}_c(t, w) = 1$ or 0 according as $\tilde{x}(t, w) < c$ or $\geqq c$. Let $e_n(s, t) = 1$ or 0 according as $t < s < t + n^{-1}$ or otherwise. Let

$$x_c^{(n)}(t, w) = \sum_j e_n(s_j, t)\tilde{x}_c(s_j, w)$$

Clearly we have

$$\{(t, w) : x^{(n)}(t, w) < c\} = \{(t, w) : x_c^{(n)}(t, w) > 0\} .$$

It is easy to see that $x_c^{(n)}(\circ, \circ)$ is measurable $\mathscr{B} \times \mathscr{F}$: hence both sets above belong to $\mathscr{B} \times \mathscr{F}$. This being true for every c, $x^{(n)}(\circ, \circ)$ is measurable $\mathscr{B} \times \mathscr{F}$ and so is $x(\circ, \circ)$.

4. Let $\{x_t, t \in T\}$ be the Markov chain described in Section 2, and let $\mathscr{F}\{x_s, 0 \leqq s \leqq t\}$ be the augmented[10] Borel field generated by the family of random variables $\{x_s, 0 \leqq s \leqq t\}$. Let $\{\mathscr{F}_t, t \in T\}$ be a non-decreasing family of Borel fields such that

$$\mathscr{F}\{x_s, 0 \leqq s \leqq t\} \subseteqq \mathscr{F}_t \subseteqq \mathscr{F}$$

and such that for every $j \in I$, $0 \leqq s \leqq t$, we have

$$\mathbf{P}\{x_t(w) = j \mid \mathscr{F}_s\} = p_{x_s, j}(t - s) \tag{3}$$

with probability one. It follows from this that we have more generally: if $\Lambda \in \mathscr{F}\{x_t, t \geqq s\}$,

$$\mathbf{P}\{\Lambda \mid \mathscr{F}_s\} = \mathbf{P}\{\Lambda \mid x_s\}$$

with probability one. Observe that if $\mathscr{F}_t = \mathscr{F}\{x_s, 0 \leqq s \leqq t\}$, the equation (3) is a consequence of the properties of the Markov chain. The introduction of larger Borel fields facilitates applications; for example, $\mathscr{F}_t$ may contain certain sets independent of those in $\mathscr{F}\{x_s, 0 \leqq$

[9] L. J. Savage showed me an example where the Borel measurability of $\tilde{x}$ does not imply that of x, without the intervention of separability.

[10] A subfield of $\mathscr{F}$ is augmented in case it contains all sets of $\mathbf{P}$-measure zero.

$s \leqq t\}$. In what follows the family of fields $\{\mathscr{F}_t, t \in T\}$ associated with the Markov chain $\{x_t, t \in T\}$ in the way described above will be fixed.

DEFINITION 2. Let $\Delta \in \mathscr{F}$. A nonnegative random variable defined and finite-valued on Δ is said to be optional (relative to $\{\mathscr{F}_t, t \in T\}$) in case for each $t \in T$, the set $\{w : \alpha(w) \leqq t\}$[11] belongs to $\mathscr{F}_t$.

We have allowed α to be undefined or infinite on $\Omega - \Delta$ in order to cover all interesting applications. We assume always that $\mathbf{P}(\Delta) > 0$; the pertinent probability triple for α is then the reduced triple $(\Delta, \Delta\mathscr{F}, \mathbf{P}(\circ|\Delta))$. If α is optional, the collection of sets Λ in $\Delta\mathscr{F}$ such that $\Lambda \cap \{\alpha(w) \leqq s\} \in \mathscr{F}_s$ for every $s \geqq 0$ is a Borel field to be denoted by $\mathscr{F}_\alpha$; when $\mathscr{F}_t = \mathscr{F}\{x_s, 0 \leqq s \leqq t\}$ it may be denoted by $\mathscr{F}\{x_t, t \leqq \alpha\}$.

DEFINITION 3. Let $\{x(t, w), t \in T\}$ be a Borel measurable Markov chain and let α be an optional random variable with domain of (definition and) finiteness Δ. Then the process $\{y(t, w), t \in T\}$ defined on the probability triple $(\Delta, \Delta\mathscr{F}, \mathbf{P}(\circ|\Delta))$ as follows:

$$y(t, w) = x(\alpha(w) + t, w), \qquad t \in T \tag{4}$$

is called the post-α process relative to the x-process.

We remark that even if x is not Borel measurable it may happen that (4) defines a process [2]; in such a case the above definition will naturally be in force. On the other hand, if x is Borel measurable, then (4) defines a process if x is not Markovian and α is not optional. Finally, the random variables of the post-α process need not be finite-valued with probability one – we are using the term "process" in a liberal sense.

The following theorem is a preliminary form of the main result of this paper. It will be completed in the Corollary to Theorem 5.

THEOREM 3. *Let $\{\hat{x}(t, w), t \in T\}$ be a separable measurable Markov chain with the state space $\overline{I}$ and the stationary transition matrix $((p_{ij}))$ satisfying the condition* (1); *and let α be an optional random variable with domain of finiteness Δ. Let $x(t, w) = \underline{\lim}_{s \downarrow t}\, x(s, w)$.*

The post-α process relative to the x-process is then a Markov chain in T^0 on $(\Delta, \Delta\mathscr{F}, \mathbf{P}(\circ|\Delta))$ with state space $J \subseteq \overline{I}$, with $\mathbf{P}\{y(t, w) = \infty\} = 0$ for every $t \in T^0$, and with the stationary transition matrix $((p_{ij}))$ restricted to J. Furthermore, if $\Lambda \in \mathscr{F}_\alpha$, $s > 0$ and

$$M_1 \in \Delta\mathscr{F}\{y_t, 0 < t < s\}, \; M_2 \in \Delta\mathscr{F}\{y_t, t \geqq s\},$$

then we have

[11] As shown by Doob, this may be relaxed to $\{w: \alpha(w) < t\}$, since (3) remains true if $\mathscr{F}_s$ is replaced by $\mathscr{F}_{s+0}$, by martingale theory.

$$(5) \qquad \mathbf{P}\{\Lambda M_1M_2 \mid y_s(w)\} = \mathbf{P}\{\Lambda M_1 \mid y_s(w)\} \cdot \mathbf{P}\{M_2 \mid y_s(w)\}$$

with probability one.

REMARK 1. The last equation is an extension of the Markovian property, which may be roughly expressed by saying that the past and the future are independent, conditioned on the present [5, p. 83]. Here in the extension the past, present, and future are, so to speak, optionally determined.

REMARK 2. There is just one missing link, namely $y_0(w)$. The post-process is defined even in T and Borel measurable, but $y_0(w)$ may assume the adjoined state ∞ with positive (even unit) probability, and transitions from there are *not* prescribed by $p_{\infty j}(\circ)$ as given in Section 2. To extend Theorem 3 to cover y_0 is the goal of the so-called theory of "boundaries." Without such a ramification we can only deal with the finite values of y_0 (see Theorem 5 and its corollary).

RESTATEMENT OF THE CONCLUSIONS IN THEOREM 3. *If* $0 < t_1 < t_2 < \cdots < t_N$, *and* $j_1, j_2, \cdots, j_N$ *are arbitrary states (inclusive of* ∞*), we have*

$$(6) \qquad \begin{aligned} &\mathbf{P}\{\Lambda : y(t_v, w) = j_v, 1 \leqq v \leqq N\} \\ &\quad = \mathbf{P}\{\Lambda ; y(t_1, w) = j_1\} \textstyle\prod_{v=1}^{N-1} p_{j_v j_{v+1}}(t_{v+1} - t_v) . \end{aligned}$$

Furthermore $\mathbf{P}\{y(t, w) = \infty\} = 0$ *for every* $t > 0$.

The complete equivalence of the two statements of Theorem 3 becomes clear if we take $s = t_n$ for a certain n, $1 \leqq n \leqq N$, and set

$$\begin{aligned} M_1 &= \{w : y(t_v, w) = j_v, 1 \leqq v \leqq n - 1\} , \\ M_2 &= \{w : y(t_v, w) = j_v, n + 1 \leqq v \leqq N\} \end{aligned}$$

5. PROOF OF THEOREM 3 IN THE FORM (6). By Theorem 2, since (2′) is satisfied, the process $\{x(t, w), t \in T\}$ is a separable, Borel measurable, standard modification of the x-process; moreover for almost all w, $x(\circ, w)$ is right lower semicontinuous in T. The separability set will be the set of terminating dyadic expansions $\{m2^{-n}, m \geqq 0, n \geqq 1\}$.

Before we proceed further it will be convenient to list the "continuity properties" needed in the sequel. They will be referred to as (i)–(iv).

(i) Each p_{ij} is uniformly continuous in T. [1 ; (2.5)].

(ii) Let

$$S_i(w) = \{t : x(t, w) = i\}, \qquad i \in \overline{I}$$

so that $\{w : x(t, w) = i\} = \{w : t \in S_i(w)\}$. Let $S_i^+(w)$, or $S_i^-(w)$, be the set of t such that $(t, t + \varepsilon) \cap S_i(w) \neq 0$, or $(t - \varepsilon, t) \cap S(w) \neq 0$, for every $\varepsilon > 0$; and $\overline{S_i(w)} = S_i(w) \cup S_i^+(w) \cup S_i^-(w)$ so that $\overline{S_i(w)}$ is the closure of

$S_i(w)$. Then for a fixed t, $i \neq \infty$, and almost all w:

$$\{w : t \in S_i(w)\} = \{w : t \in \overline{S_i(w)}\} = \{w : t \in S_i^+(w)\} = \{w : t \in S_i^-(w)\} .$$

(The last equation is to be omitted if $t = 0$.) [2 ; Theorem 2].

(iii) For almost all w, the sample function $x(\circ, w)$ has the following property: for every t, as $s \uparrow t$ or $s \downarrow t$, $x(s, w)$ has at most one finite limiting value, [3, Theorem 12] or [1, Theorem 4].

(iv) For almost all w, the sample function $x(\circ, w)$ has the following property: for every $i \neq \infty$, $S_i^+(w) \cap S_i^-(w) \subseteqq S_i(w) \subseteqq S_i^+(w) \cup S_i^-(w)$, [1 ; Theorem 4].[12]

In the sequel we may suppose, without loss of generality, that the sample function properties given in (ii)-(iv) are true for all w.

We now define a w-function γ_j as follows:

$$(7) \qquad \gamma_j(w) = \inf\{t : t > \alpha(w), x(t, w) = j\} \qquad j \in I$$

From the second part of (iv) we know that for almost all w this infimum must be a limit point from the right of $S_j(w)$. The domain of (definition and) finiteness of γ_j is

$$\Gamma_j = \Delta \cap \{w : S_j(w) \cap [\alpha(w), \infty) \neq 0\} .$$

For every $t > 0$, we have

$$\Gamma_j \cap \{w : \gamma_j(w) - \alpha(w) \leqq t\}$$
$$= \Gamma_j \cap_{m=1}^{\infty} \{w : S_j(w) \cap [\alpha(w), \alpha(w) + t + m^{-1}] \neq 0\} .$$

Let $f(s, w) = 1$ if $x(s, w) = j$ and 0 otherwise; let $g(s, w) = 1$ if $\alpha(w) \leqq s \leqq \alpha(w) + t + m^{-1}$ and 0 otherwise. Both f and g are measurable $\mathscr{B} \times \mathscr{F}$. Let $\{s_k\}$ be a separability set for the x-process, then $h(\circ) = \sum_k f(s_k, \circ)g(s_k, \circ)$ is measurable $\mathscr{F}$. We have, up to a set of probability zero,

$$\Gamma_j \cap \{w : S_j(w) \cap [\alpha(w), \alpha(w) + t + m^{-1}] \neq 0\} = \Gamma_j \cap \{w : h(w) > 0\} .$$

Hence γ_j is a random variable[13] and it follows from the optionality of α that for every t,

$$(8) \qquad \{w : \gamma_j(w) \leqq t\} \in \cap_{m=1}^{\infty} \mathscr{F}_{t+m^{-1}} .$$

For any set $\mathscr{X} \in \mathscr{F}$ with $\mathbf{P}(\mathscr{X}) > 0$, and any random variable ξ, the conditional probability $\mathbf{P}\{\cdots | \mathscr{X} ; \xi\}$ is simply the ordinary conditional probability with respect to ξ of the basic probability measure $\mathbf{P}\{\cdots | \mathscr{X}\}$ on the couple $(\mathscr{X}, \mathscr{X}\mathscr{F})$. We use similar notations for

[12] We remark that in [1] and [2] the process is assumed only to be measurable, not necessarily Borel measurable.

[13] According to the new definition given in the footnote to Definition 2, γ_j is optional.

the mathematical expectation $\mathbf{E}$; and also write $\int\{\xi \| \mathscr{X}\}$ for the integral of ξ over $\mathscr{X}$ with respect to the measure $\mathbf{P}$.

Let $\Lambda \in \mathscr{F}_\alpha$. Let

$$A(\Lambda\,;\, s) = \mathbf{P}\{\Lambda\,;\, \alpha(w) \leqq s\}\ ,$$

and call the corresponding measure the $A(\Lambda, \circ)$ measure. In the sequel where the range of α is in question the phrase "almost all," abbreviated to "a. a.," shall mean "all except a set of $A(\Lambda, \circ)$ measure zero." Defien a conditional probability distribution in the wide sense of γ_j with respect to Λ and α, to be denoted by $C_j(s, B|\Lambda)$, where $s \in T$ and $B \in \mathscr{B}$, as follows: (a) for each s, $C_j(s, \circ|\Lambda)$ is a measure on $\mathscr{B}$; (b) for each B, $C_j(\circ, B|\Lambda)$ is a Baire function on T; and (c) for every $B_1 \in \mathscr{B}$, $B_2 \in \mathscr{B}$ we have

$$\int_{B_1} C_j(s, B_2|\Lambda) d_s A(\Lambda\,;\, s) = \mathbf{P}\{\Lambda\,;\, \alpha(w) \in B_1\,;\, \gamma_j(w) \in B_2\}\ .$$

Such a choice of C_j is possible [5, p. 29]. Thus

$$C_j(s, B|\Lambda) = \mathbf{P}\{\gamma_j(w) \in B|\Lambda\,;\, \alpha(w) = s\}$$

for a. a. s. We shall write $C_j(s, u|\Lambda)$ for $C_j(s, B|\Lambda)$ if $B = [0, u]$.

6. The following lemma is basic.

Lemma 1. *If $t \geqq 0$ and $j \in I$, we have almost everywhere on $\{w\colon \gamma_j(w) \leqq t\}$*

$$\mathbf{P}\{x(t, w) = j|\Lambda\,;\, \alpha(w), \gamma_j(w)\} = p_{jj}(t - \gamma_j(w))$$

Proof. Define[14] for each $n \geq 0$,

$$\gamma_j^{(n)}(w) = \min_{m \geq 1}\{m2^{-n} : \gamma_j(w) < m2^{-n}\,;\, x(m2^{-n}, w) = j\}\ .$$

Clearly $\gamma_j^{(n)}$ is a random variable and $\gamma_j^{(n)}(w) \downarrow \gamma_j(w)$ for almost all w in Γ_j. According to (ii), $\mathbf{P}\{\alpha(w) < \gamma_j(w) = t\} = 0$ for every t. Hence we have if $s \leqq s' < t$,

$$\begin{aligned}
&\mathbf{P}\{\Lambda\,;\, \alpha(w) \leqq s\,;\, \gamma_j(w) \leqq s'\,;\, x(t, w) = j\} \\
&\quad = \lim_{n\to\infty} \mathbf{P}\{\Lambda\,;\, \alpha(w) \leqq s\,;\, \gamma_j(w) \leqq s'\,;\, \gamma_j^{(n)}(w) < t\,;\, x(t, w) = j\} \\
&\quad = \lim_{n\to\infty} \textstyle\sum_{m < t2^n} \mathbf{P}\{\Lambda\,;\, \alpha(w) \leqq s\,;\, \gamma_j(w) \leqq s'\,;\, \gamma_j^{(n)}(w) \\
&\quad = m2^{-n}\,;\, x(t, w) = j\}\ .
\end{aligned}$$

Since α is optional and γ_j has the property (8), the set

$$\Lambda \cap \{w : \alpha(w) \leqq s\,;\, \gamma_j(w) \leqq s'\,;\, \gamma_j^{(n)}(w) = m2^{-n}\}$$

belongs to $\mathscr{F}_{m2^{-n}}$. Hence we have, using the properties of $\{\mathscr{F}_t, t \in T\}$,

---

[14] The introduction of $\gamma_j^{(n)}$ simplifies a lot of the formularismus. I take this opportunity to correct similar arguments on p. 39 of [1] and p. 197 of [2] by using this device.

$$\begin{aligned}
&\mathbf{P}\{\Lambda\,;\,\alpha(w)\leqq s\,;\,\gamma_j(w)\leqq s'\,;\,x(t,w)=j\}\\
&\quad=\lim_{n\to\infty}\textstyle\sum_{m<t2^n}\mathbf{P}\{\Lambda\,;\,\alpha(w)\leqq s;\,\gamma_j(w)\leqq s',\,\gamma_j^{(n)}(w)=m2^{-n}\}\,p_{jj}(t-m2^{-n})\\
&\quad=\lim_{n\to\infty}\int\{p_{jj}(t-\gamma_j^{(n)}(w))\,\|\,\Lambda\,;\,\alpha(w)\leqq s\,;\,\gamma_j(w)\leqq s'\,;\,\gamma_j^{(n)}(w)<t\}\\
&\quad=\int\{p_{jj}(t-\gamma_j(w))\,\|\,\Lambda\,;\,\alpha(w)\leqq s\,;\,\gamma_j(w)\leqq s'\}
\end{aligned}$$

Furthermore, we have by (iii) and the right lower semi-continuity of $x(\circ, w)$,

$$\begin{aligned}
\mathbf{P}\{\Lambda\,;\,\alpha(w)=t\,;\,\gamma_j(w)\leqq t\,;\,x(t,w)=j\} &= \mathbf{P}\{\Lambda\,;\,\alpha(w)=\gamma_j(w)=t\}\\
&= \int\{p_{jj}(t-\gamma_j(w)\,\|\,\Lambda\,;\,\alpha(w)=\gamma_j(w)=t\}\ .
\end{aligned}$$

Combining these two cases we see that the equation

$$\begin{aligned}
&\mathbf{P}\{\Lambda\,;\,\alpha(w)\leqq s\,;\,\gamma_j(w)\leqq s'\,;\,x(t,w)=j\}\\
&\qquad= \int\{p_{jj}(t-\gamma_j(w))\,\|\,\Lambda\,;\,\alpha(w)\leqq s\,;\,\gamma_j(w)\leqq s'\}
\end{aligned}$$

is valid if $0\leqq s\leqq s'\leqq t$. This is equivalent to the assertion of Lemma 1.

LEMMA 2. *One version of the conditional probability*

$$\mathbf{P}\{x(t,w)=j\,|\,\Lambda\,;\,\alpha(w)=s\} \tag{9}$$

is given by

$$r_j(s,t\,|\,\Lambda)=\int_{[s,t]}p_{jj}(t-u)d_uC_j(s,u\,|\,\Lambda)\,,\qquad j\in I,\ 0\leqq s\leqq t. \tag{10}$$

For each $s\geqq 0$, $r_j(s,\circ\,|\,\Lambda)$ *is right continuous in* $[s,\infty)$, *and* $r_j(\circ,\circ\,|\,\Lambda)$ *is a Baire function of the pair* (s,t) *in* $0\leqq s\leqq t$.

PROOF. We have with probability one

$$\mathbf{P}\{x(t,w)=j\,|\,\Lambda\,;\,\alpha(w)\}=\mathbf{E}\{\mathbf{P}[x(t,w)=j\,|\,\Lambda\,;\,\alpha(w),\,\gamma_j(w)]\,|\,\Lambda\,;\,\alpha(w)\}\ .$$

Hence using Lemma 1 and the C_j specified above, the right member of (10) is one version of (9). It is clearly right continuous in t, and is a Baire function of (s,t) by (i) and the properties of C_j, being the limit of Riemann-Stieltjes sums of Baire functions.

LEMMA 3. *We have*

$$r_k(s,t+t'\,|\,\Lambda)=\textstyle\sum_j r_j(s,t\,|\,\Lambda)p_{jk}(t'),\quad \text{a. a. } s,\ t>s,\ t'\geqq 0\,; \tag{11}$$

and

$$\textstyle\sum_j r_j(s,t\,|\,\Lambda)=1 \qquad\qquad \text{a. a. } s,\ t>s, \tag{12}$$

where the summations are over I. *In particular, for* a. a. s, $r_j(s,\circ\,|\,\Lambda)$ *is continuous in* (s,∞).

PROOF. Since α is optional, we have for every $k\in I$,

$$\mathbf{P}\{\Lambda\,;\,\alpha(w)\leqq s\,;\,x(t+t',w)=k\}=\textstyle\sum_j P\{\Lambda\,;\,\alpha(w)\leqq s\,;\,x(t,w)=j\}\,p_{jk}(t').$$

It follows from the defining property of a conditional probability that

$$\mathbf{P}\{x(t+t', w) = k \mid \Lambda\,;\, \alpha(w)\} = \sum_j \mathbf{P}\{x(t, w) = j \mid \Lambda\,;\, \alpha(w)\}\, p_{jk}(t')$$

for almost all w ; that is

$$r_k(s, t+t' \mid \Lambda) = \sum_j r_j(s, t \mid \Lambda) p_{jk}(t'), \quad t, t' \geqq 0, \text{ a. a. } s \text{ in}[0, t].^{15}$$

Both sides being Baire functions of (s, t, t'), we can use Fubini's theorem to replace the last proviso by a. a. s, a. a. (t, t') in $[s, \infty) \times [0, \infty)$. Here and hereafter a. a. t or (t, t') refer to the Lebesgue measure in one or two dimensions. Using the right continuity and Fatou's lemma, we see that

$$(13) \quad r_k(s, t+t' \mid \Lambda) \geqq \sum_j r_j(s, t \mid \Lambda) p_{jk}(t') \quad \text{a. a. } s, \ (t, t') \in [s, \infty) \times (0, \infty).$$

Summing over k, we have

$$(14) \qquad \sum_k r_k(s, t+t' \mid \Lambda) \geqq \sum_j r_j(s, t \mid \Lambda)\,.$$

Since

$$\sum_j \int_{[0,t]} r_j(s, t \mid \Lambda) dA(\Lambda\,;\, s) = \sum_j \mathbf{P}\{\Lambda\,;\, \alpha(w) \leqq t\,;\, x(t, w) = j\}$$
$$= \mathbf{P}\{\Lambda\,;\, \alpha(w) \leqq t\}\ ,$$

it follows from the defining property of a conditional probability the

$$\sum_j r_j(s, t \mid \Lambda) = 1 \qquad t \geqq 0, \text{ a. a. } s \text{ in } [0, t].$$

Using Fubini's theorem again, the last proviso may be replaced by a. a. s, a. a. t in $[s, \infty)$. This together with (14) gives (12). Hence there is equality also in (13), namely (11) is true. The continuity of $r_j(s, \circ \mid \Lambda)$ in (s, ∞), for a. a. s, now follows from (11), (12) and (i).

7. We are now ready to resume the proof of Theorem 3. Without loss of generality we may suppose that Ω has been so chosen that the sample function property asserted in (ii)-(iv) is true for all w. Consequently for any random variable $\tau(w) \geqq 0$ we have

$$\bigcap_m \left\{w : S_i(w) \cap \left[\tau(w) - \frac{1}{m}, \tau(w)\right] \neq 0\,;\, S_i(w) \cap \left[\tau(w), \tau(w) + \frac{1}{m}\right] \neq 0\right\}$$
$$\subseteqq \{w : x(\tau(w), w) = i\} \subseteqq \bigcap_m \left\{w : S_i(w) \cap \left[\tau(w) - \frac{1}{m}, \tau(w) + \frac{1}{m}\right] \neq 0\right\}.$$

(The second inclusion is trivial.) Thus if $0 < t_1 < \cdots < t_N$, and $j_1, \cdots, j_N$ are arbitrary states $\neq \infty$ we have

[15] The set of measure zero involved in "a. a." written after other variables may depend on these variables; e.g., here the exceptional set of s in $[0, t]$ may depend on t and t'.

$$\Lambda \cap_m \cap_{v=1}^{N}\left\{w : S_{j_v}(w) \cap \left[\alpha(w) + t_v - \frac{1}{m}, \alpha(w) + t_v\right] \neq 0 ;\right.$$

$$\left. S_{j_v}(w) \cap \left[\alpha(w) + t_v, \alpha(w) + t_v + \frac{1}{m}\right] \neq 0\right\}$$

$$\subseteqq \Lambda \cap_{v=1}^{N}\{w : x(\alpha(w) + t_v, w) = j_v\}$$

$$\subseteqq \Lambda \cap_m \cap_{v=1}^{N}\left\{w : S_{j_v}(w) \cap \left[\alpha(w) + t_v - \frac{1}{m}, \alpha(w) + t_v + \frac{1}{m}\right] \neq 0\right\}.$$

Call these three w-sets Λ_1, Λ_2 and Λ_3 in order of their appearance. It is seen that Λ_1 and Λ_3 both belong to $\mathscr{F}$, by an argument similar to the previous one showing that γ_j is a random variable. Since $x(\alpha(\circ) + t_v, \circ)$ is measurable $\mathscr{F}$ as noted before, $\Lambda_2 \in \mathscr{F}$.

Next, we are going to prove that

$$\mathbf{P}(\Lambda_3) \leqq Q \leqq \mathbf{P}(\Lambda_1)$$

where Q is to be specified later. It will follow from this that all the three sets Λ_1, Λ_2 and Λ_3 have probability equal to Q. This will establish Theorem 3 in the form (6).

Consider the sets :

$$A_{h,m} = \Lambda \cup_{n=0}^{\infty} \left\{w : nh \leqq \alpha(w) < (n+1)h ;\right.$$

$$\left. S_{j_v}(w) \cap \left[t_v + nh - \frac{1}{m}, t_v + (n+1)h + \frac{1}{m}\right] \neq 0, 1 \leqq v \leqq N\right\}$$

$$B_m = \Lambda \cap \left\{w : S_{j_v}(w) \cap \left[\alpha(w) + t_v - \frac{1}{m}, \alpha(w) + t_v + \frac{1}{m}\right] \neq 0, 1 \leqq v \leqq N\right\},$$

where $h = 2^{-p}$ and henceforth $h \to 0$ means $p \to \infty$. It is clear that $A_{2h,m} \supseteqq A_{h,m}$ so that $\cap_h A_{h,m}$ exists ; in fact it is equal to B_m. Hence $\mathbf{P}(B_m) = \lim_{h\to 0} \mathbf{P}(A_{h,m})$, and

$$\mathbf{P}(\Lambda_3) = \lim_{m\to\infty} \lim_{h\to 0} \mathbf{P}(A_{h,m})$$

where the repeated limit obviously exists.

Now choose an arbitrary s in $(0, t_1)$. We have then for $t_1 - s > h + \frac{1}{m}$,

$$\mathbf{P}(A_{h,m})$$

$$= \sum_{n=0}^{\infty} \sum_j \mathbf{P}\{\Lambda ; nh \leqq \alpha(w) < (n+1)h ; x((n+1)h + s, w) = j\} \cdot$$

$$\mathbf{P}\left\{S_{j_v}(w) \cap \left[t_v + nh - \frac{1}{m}, t_v + (n+1)h + \frac{1}{m}\right] \neq 0,\right.$$

$$\left. 1 \leqq v \leqq N \,|\, x((n+1)h + s, w) = j\right\}$$

$$= \sum_{n=0}^{\infty} \sum_{j} b(n, h, j)\, \mathbf{P}\left\{S_{j_v}(w) \cap \left[t_v - s - h - \frac{1}{m}, t_v - s + \frac{1}{m}\right] \neq 0, \; 1 \leqq v \leqq N \mid x(0, w) = j\right\}$$

where

$$b(n, h, j) = \mathbf{P}\{\Lambda \,;\, nh \leqq \alpha(w) < (n+1)h \,;\, x((n+1)h + s, w) = j\} .$$

Clearly for each m,

$$\lim_{h\to 0} \mathbf{P}(A_{h,m}) \leqq \lim_{h\to 0} \sum_{n=0}^{\infty} \sum_{j} b(n, h, j)\mathbf{P}\left\{S_{j_v}(w) \cap \left[t_v - s - \frac{2}{m}, t_v - s + \frac{1}{m}\right] \neq 0, \; 1 \leqq v \leqq N \mid x(0, w) = j\right\}$$

$$= \lim_{h\to 0} Q_{h,m}$$

say. The last limit is easily seen to exist, in fact it is equal to

$$\mathbf{P}\left\{\Lambda \,;\, S_{j_v}(w) \cap \left[\alpha(w) + t_v - \frac{2}{m}, \alpha(w) + t_v + \frac{1}{m}\right] \neq 0, 1 \leqq v \leqq N\right\} .$$

We now evaluate

$$\lim_{m\to\infty} \lim_{h\to 0} Q_{h,m} \tag{15}$$

by inverting the repeated limit. It follows from (ii) that

$$\lim_{m\to\infty} \mathbf{P}\left\{S_{j_v}(w) \cap \left[t_v - s - \frac{2}{m}, t_v - s + \frac{1}{m}\right] \neq 0, \; 1 \leqq v \leqq N \mid x(0, w) = j\right\} \tag{16}$$
$$= \mathbf{P}\{x(t_v - s, w) = j_v, 1 \leqq v \leqq N \mid x(0, w) = j\} = p_{jj_1}(t_1 - s)R ,$$

where, and henceforth, we use the abbreviation

$$R = \prod_{v=1}^{N-1} p_{j_v j_{v+1}}(t_{v+1} - t_v) .$$

Consequently

$$\lim_{h\to 0} \lim_{m\to\infty} Q_{h,m} = \lim_{h\to 0} \sum_{n=0}^{\infty} \sum_{j} b(n, h, j) p_{jj_1}(t_1 - s)R . \tag{17}$$

To show that the last limit exists, we use the r_j introduced above and write

$$b(n, h, j) = \int_{nh}^{(n+1)h} r_j(u, (n+1)h + s \mid \Lambda) dA(\Lambda \,;\, u) . \tag{18}$$

The limit (17) is thus equal to

$$\lim_{h\to 0}\sum_{n=0}^{\infty}\sum_j\int_{nh}^{(n+1)h} r_j(u,\ (n+1)h+s\,|\,\Lambda)p_{jj_1}(t_1-s)R\,dA(\Lambda\ ;\ u)$$
$$=\lim_{h\to 0}\sum_{n=0}^{\infty}\int_{nh}^{(n+1)h} r_{j_1}(u,\ (n+1)h+t_1\,|\,\Lambda)R\,dA(\Lambda\ ;\ u)$$

on account of Lemma 3 ; and hence equal to

$$\lim_{h\to 0}\int_0^\infty r_{j_1}\Big(u,\left[\frac{u}{h}\right]h+h+t_1\,|\,\Lambda\Big)R\,dA(\Lambda\ ;\ u)$$
$$=\int_0^\infty r_{j_1}(u,\ u+t_1\,|\,\Lambda)R\,dA(\Lambda\ ;\ u)=Q\ .$$

The passage of limit under the integral is justified since $0\leqq r_{j_1}\leqq 1$ and for each u, $r_{j_1}(u,\ \circ\,|\,\Lambda)$ is continuous in $[u,\ \infty)$ by Lemma 3.

It remains to justify the inversion of the limit in (15). This will be done by proving that the convergence of $Q_{h,m}$ as $m\to\infty$ is uniform in h. Now $Q_{h,m}$ is of the form

$$\sum_{n=0}^{\infty}\sum_j b(h,\ n,\ j)c(j,\ m) \tag{19}$$

where only the relevant parameters have been indicated, and

$$0\leqq c(j,\ m)\leqq 1$$
$$\lim_{m\to\infty} c(j,\ m)=c(j,\ \infty)$$

for each j. Since

$$\sum_{n=0}^{\infty}\sum_j b(h,\ n,\ j)=\mathbf{P}(\Lambda)$$

there is dominated convergence in (19) with respect to m. Since $c(j,\ m)$ involves j but not n, it is sufficient, in order to prove the asserted uniform convergence, to prove that the series $\sum_j\sum_{n=0}^{\infty} b(n,\ h,\ j)$ converges in j uniformly with respect to h. Our previous calculation following (18) yields

$$\sum_j\sum_{n=0}^{\infty} b(h,\ n,\ j)=\sum_j\int_0^\infty r_j\Big(u,\left[\frac{u}{h}\right]h+h+s\,|\,\Lambda\Big)dA(\Lambda\ ;\ u)\ .$$

Let us put

$$c_j(h)=\int_0^\infty r_j\Big(u,\left[\frac{u}{h}\right]h+h+s\,|\,\Lambda\Big)dA(\Lambda\ ;\ u)\ ,\qquad h>0;$$
$$c_j(0)=\int_0^\infty r_j(u,\ u+s\,|\,\Lambda)dA(\Lambda\ ;\ u)\ .$$

By Lemma 3, $r_j(u,\ [uh^{-1}]h+h+s\,|\,\Lambda)$ as a function of h is continuous at h except when u is an integral multiple of h, in which case it is right continuous at h but the left limit is $r_j(u,\ [uh^{-1}]h+s\,|\,\Lambda)$. It follows that

for each $h > 0$, c_j has left and right limits $c_j(h-)$ and $c_j(h+)$. Moreover it follows from (12) of Lemma 3 that

$$\begin{aligned}\sum_j c_j(h) &= \sum_j c_j(h+) = \sum_j c_j(h-)\\ &= \int_0^\infty \sum_j r_j(u, [uh^{-1}]h + \Theta(u)h + s\,|\,\Lambda)dA(\Lambda\,;u) = \int_0^\infty dA(\Lambda\,;u) = \mathbf{P}(\Lambda)\end{aligned}$$

where $\Theta(u) = 0$ or 1. Finally $c_j(0+) = c_j(0)$ and

$$\sum_j c_j(0) = \sum_j c_j(0+) = \mathbf{P}(\Delta)\,.$$

We now invoke the following generalization of Dini's theorem[16] to prove the uniform convegence of $\sum_j c_j(h)$ with respect to h in $[0, 1]$.

GENERALIZATION OF DINI'S THEOREM. Let $\{c_j, j \geqq 1\}$ be a sequence of non-negative functions in a closed interval C, such that $\sum_j c_j(h) = c(h)$ if $h \in C$, where c is a finite function in C. Suppose that each of the functions c and c_j has left and right limits at each h interior to C and a right (left) limit at the left (right) endpoint of C such that

$$\sum_j c_j(h-) = c(h-)\,, \qquad \sum_j c_j(h+) = c(h+)\,.$$

Then the series $\sum_j c_j(h) = c(h)$ converges uniformly with respect to h in C.

This is proved in the same way as Dini's theorem, e.g., [12; p. 13] and withal the evaluation of $\mathbf{P}(\Lambda_3)$ is concluded.

The evaluation of $\mathbf{P}(\Lambda_1)$ is much simpler. We have

$$\begin{aligned}\mathbf{P}(\Lambda_1) &= \lim_{m\to\infty}\lim_{h\to 0}\sum_{n=0}^\infty \sum_j b(h, n, j)\mathbf{P}\Big\{S_{j_v}(w) \cap \Big[t_v + nh - \frac{1}{m},\\ &\qquad t_v + (n+1)h\Big] \neq 0\,;\, S_{j_v}(w) \cap \Big[t_v + nh,\, t_v + (n+1)h + \frac{1}{m}\Big] \neq 0\,;\\ &\qquad\qquad 1 \leqq v \leqq N\,|\,x((n+1)h + s, w) = j\Big\}\\ &= \lim_{m\to\infty}\lim_{h\to 0}\sum_{n=0}^\infty \sum_j b(h, n, j)\mathbf{P}\Big\{S_{j_v}(w) \cap \Big[t_v - s - h - \frac{1}{m},\\ &\qquad t_v - s\Big] \neq 0\,;\, S_{j_v}(w) \cap \Big[t_v - s - h,\, t_v - s + \frac{1}{m}\Big] \neq 0\,;\\ &\qquad\qquad 1 \leqq v \leqq N\,|\,x(0, w) = j\Big\}\\ &\geqq \lim_{m\to\infty}\lim_{h\to 0}\sum_{n=0}^\infty \sum_j b(h, n, j)\mathbf{P}\{x(t_v - s, w) = j_v,\\ &\qquad\qquad 1 \leqq v \leqq N\,|\,x(0, w) = j\}\end{aligned}$$

[16] Instead of the generalization we may use Dini's theorem itself on the continuous part of $A(\Lambda\,;\,\circ)$ and treat the purely discontinuous part in a trivial way.

$$= \lim_{h\to 0} \sum_{n=0}^{\infty} \sum_j b(h, n, j) p_{jj_1}(t_1 - s) R$$

$$= \int_0^\infty r_{j_1}(u, u + t_1 | \Lambda) R dA(\Lambda ; u) = Q .$$

We have therefore proved that

$$\text{(20)} \quad \mathbf{P}\{\Lambda ; y(t_v, w) = j_v, 1 \leqq v \leqq N\} = Q = \int_0^\infty r_{j_1}(u, u + t_1 | \Lambda) R dA(\Lambda ; u) .$$

In particular if $N = 1$ we have

$$\text{(21)} \qquad \mathbf{P}\{\Lambda ; y(t, w) = j\} = \int_0^\infty r_j(u, u + t | \Lambda) dA(\Lambda ; u) .$$

Substituting back into (20) we obtain (6). Theorem 3 is completely proved except for Corollary 2 which follows.

COROLLARY 1. *We have*

$$\mathbf{P}\{x(\alpha(w) + t_v, w) = j_v, 1 \leqq v \leqq N | \Lambda ; \alpha(w)\} = r_{j_1}(\alpha(w), \alpha(w) + t_1 | \Lambda) R$$

with probability one.

PROOF. Replacing Λ by $\Lambda_1 = \Lambda \cap \{w : \alpha(w) \leqq s\}$ in (20) as we may, and recalling the definition of A after this substitution, we have

$$\begin{aligned} \mathbf{P}\{\Lambda ; \alpha(w) \leqq s ; y(t_v, w) = j_v, 1 \leqq v \leqq N\} \\ = \int_0^\infty r_j(u, u + t | \Lambda_1) R dA(\Lambda_1 ; u) \\ = \int \{r_j(\alpha(w), \alpha(w) + t | \Lambda) R \| \Lambda ; \alpha(w) \leqq s\} . \end{aligned}$$

Corollary 1 follows.

COROLLARY 2. *We have for every* $t > 0$,

$$\mathbf{P}\{y(t, w) = \infty | \Delta\} = 0 .$$

PROOF. Taking $\Lambda = \Delta$ in (21) and summing over j, we have

$$\begin{aligned} \mathbf{P}\{\Delta ; y(t, w) \in I\} &= \int_0^\infty \sum_{j \in I} r_j(u, u + t | \Delta) dA(\Delta ; u) \\ &= A(\Delta ; \infty) = \mathbf{P}(\Delta) . \end{aligned}$$

Corollary 2 follows.

COROLLARY 3. *For each* $t > 0$,

$$p \lim_{s\to t} y(s, \circ) = y(t, \circ)$$

where the limit in probability is relative to $\mathbf{P}(\circ | \Delta)$.

PROOF. We have by Fatou's lemma, Lemma 3 and the theorem, if $\varepsilon < 1$,

$$\begin{aligned}\lim_{s\uparrow t} \mathbf{P}\{|y(s, w) - y(t, w)| < \varepsilon\} &= \lim_{s\uparrow t} \textstyle\sum_{j\in I} \mathbf{P}\{y(s, w) = y(t, w) = j\} \\ &= \lim_{s\uparrow t} \textstyle\sum_{j\in I} \displaystyle\int_0^\infty r_j(u, u + s|\Delta)p_{jj}(t - s)dA(\Delta ; u) \\ &\geqq \textstyle\sum_{j\in I} \displaystyle\int_0^\infty \lim_{s\uparrow t} r_j(u, u + s|\Delta)p_{jj}(t - s)dA(\Delta ; u) \\ &= \textstyle\sum_{j\in I} \displaystyle\int_0^\infty r_j(u, u + t|\Delta)\, dA(\Delta ; u) = A(\Delta ; \infty) = \mathbf{P}(\Delta) .\end{aligned}$$

Similarly for $s \downarrow t$.

COROLLARY 4. *There is a standard modification of $\{y(t, w), t \in T^\circ\}$ that is separable and measurable, for which any denumerable set dense in T° is a separability set.*

This follows from Corollary 1 and a general theorem of Doob [5 ; Theorem II. 2.6].

8. Turning to the "initial distribution" of the y-process, we have for almost all $w \in \Delta$,

$$y(0, w) = \underline{\lim}_{t\downarrow 0}\, y(t, w) , \tag{22}$$

since $x(\alpha(w), w) = \underline{\lim}_{t\downarrow \alpha(w)}\, x(t, w)$ by the last assertion in Theorem 2. It follows by separability that $\mathscr{F}\{y_t, 0 \leqq t \leqq s\} = \mathscr{F}\{y_t, 0 < t \leqq s\}$ for every $s > 0$.

We define

$$r_j(\Lambda ; t) = \int_0^\infty r_j(s, s + t|\Lambda)dA(\Lambda ; s) , \qquad j \in I ; t > 0$$

and

$$r_j(\Lambda ; 0) = \lim_{t\to 0} r_j(\Lambda ; t) = \int_0^\infty r_j(s, s|\Lambda)dA(\Lambda ; s)$$

the limit existing by (10). According to (21), for each $t > 0$ the set $\{r_j(\Delta ; t), j \in I\}$ is the absolute distribution of y_t, to which we may add $r_\infty(\Delta ; t) \equiv 0$, $t > 0$. Clearly we have

$$\begin{aligned} &r_j(\Lambda ; t) \geqq 0 ; \qquad \textstyle\sum_j r_j(\Lambda ; t) = \mathbf{P}(\Lambda) && t > 0 ; \\ &r_k(\Lambda ; s + t) = \textstyle\sum_j r_j(\Lambda ; s)p_{jk}(t) && s > 0, t \geqq 0. \end{aligned}$$

The situation is different at $t = 0$.

THEOREM 4. *We have*

$$\mathbf{P}(\Lambda ; y(0, w) = j) = r_j(\Lambda ; 0) , \qquad j \in I ; \tag{23}$$

$$\mathbf{P}(\Lambda ; y(0, w) = \infty) = \mathbf{P}(\Lambda) - \textstyle\sum_{j\in I} r_j(\Lambda ; 0) \tag{24}$$

PROOF. If $\alpha(w) = \gamma_j(w)$ where γ_j is defined in (7), then by (iii) j is the

only finite limiting value of $x(t, w)$ as $t \downarrow \alpha(w)$ and hence $y(0, w)$ is equal to j by (22). Conversely if $y(0, w) = j$ then $\alpha(w) = \gamma_j(w)$ by the definition (7). We have thus

$$\{w : y(0, w) = j\} = \{w : \alpha(w) = \gamma_j(w)\} .$$

Consequently we have

$$\begin{aligned}
\mathbf{P}\{\Lambda ; y(0, w) = j\} &= \mathbf{P}\{\Lambda ; \alpha(w) = \gamma_j(w)\} \\
&= \lim_{n\to\infty}\sum_{m=1}^{\infty} \int_{[mn^{-1},(m+1)n^{-1})} C_j(s, (m+1)n^{-1}|\Lambda)dA(\Lambda ; s) \\
&= \lim_{n\to\infty}\int_0^{\infty} C_j(s, [ns+1]n^{-1}|\Lambda)dA(\Lambda ; s) \\
&= \int_0^{\infty} C_j(s, s|\Lambda)dA(\Lambda ; s).
\end{aligned}$$

But we have $C_j(s, s|\Lambda) = r_j(s, s|\Lambda)$ from (10) of Lemma 2, from which (23) follows ; (24) then follows since $y(0, w) = \infty$ if and only if $y(0, w) \notin I$.

THEOERM 5. *A necessary and sufficient condition that $\{y_t, t \in T\}$ be a stochastic process whose random variables are finite with probability one on $(\Delta, \Delta\mathscr{F}, P(\circ|\Delta))$ is that $\sum_{j\in I} r_j(\Delta ; 0) = \mathbf{P}(\Delta)$. It is then a Markov chain with state space $J \subseteqq I$, the initial distribution $\{r_j(\Delta ; 0) ; j \in I\}$ and the stationary transition matrix $((p_{ij}))$ restricted to J. Furthermore* (5) *is then true even if $s = 0$.*

PROOF. The first assertion is an immediate consequence of Theorem 4 with $\Lambda = \Delta$. Now suppose that $\sum_{j\in I} r_j(\Delta ; 0) = \mathbf{P}(\Delta)$, namely that $\sum_{j\in I} \mathbf{P}(y(0, w) = j) = \mathbf{P}(\Delta)$. In proving (5) we may prove it for any standard modification of the y-process, since the Borel fields involved are all augmented by our agreement. We shall use the version given by Corollary 4 to Theorem 3, and prove (5) in the form (6) but with the additional parameter value $t = 0$. Define for each $n \geqq 1$,

$$\delta_j^{(n)}(w) = \min_{m\geq 1}(m2^{-n} : y(m2^{-n}, w) = j) ;$$

compare with the $\gamma_j^{(n)}$ in the proof of Lemma 1. On the set

$$\{w : y(0, w) = j\} = \{w : \alpha(w) = \gamma_j(w)\}$$

we know that $\delta_j^{(n)}$ is a finite random variable and that $\delta_j^{(n)}(w) \downarrow 0$ as $n\to\infty$ (with probability one). Let

$$0 < t_1 < \cdots < t_N ; \qquad j \in I ; j_1 , \cdots , j_N \in \overline{I}$$

We have using Theorem 3,

$$\mathbf{P}\{\Lambda ; y(0, w) = j ; y(t_v, w) = j_v, 1 \leqq v \leqq N\}$$

$$= \lim_{n\to\infty}\sum_{m=1}^{\infty} \mathrm{P}\{\Lambda\ ;\ y(0,\ w)=j\ ;\ \delta_j^{(n)}(w)=m2^{-n}\ ;\ y(m2^{-n},\ w)=j\ ;\ y(t_v,\ w)=j_v,\ 1\leqq v\leqq N\}$$

$$= \lim_{n\to\infty}\sum_{m=1}^{\infty} \mathrm{P}\{\Lambda\ ;\ y(0,\ w)=j\ ;\ \delta_j^{(n)}(w)=m2^{-n}\ ;\ y(m2^{-n},\ w)=j\}p_{jj_1}(t-m2^{-n})R$$

$$= \lim_{n\to\infty}\sum_{m=1}^{\infty}\int\{p_{jj_1}(t-\delta_j^{(n)}(w))R\,\|\,\Lambda\ ;\ y(0,w)=j\ ;\ \delta_j^{(n)}(w)=m2^{-n}\}.$$

Observe that

$$\{w : y(0,\ w)=j\ ;\ \delta_j^{(n)}(w)=m2^{-n}\}\ \in\ \mathscr{F}\{y_t,\ 0<t\leqq m2^{-n}\}$$

The above is equal to

$$\lim_{n\to\infty}\int\{p_{jj_1}(t-\delta_j^{(n)}(w))R\,\|\,\Lambda\ ;\ y(0,\ w)=j\} = \int\{p_{jj_1}(t)R\,\|\,\Lambda\ ;\ y(0,\ w)=j\}$$
$$= \mathrm{P}\{\Lambda\ ;\ y(0,\ w)=j\}p_{jj_1}(t)R\ .$$

This is the extension of (6) to include $t=0$ and it establishes Theorem 5 ; indeed we have proved a little more :

COROLLARY. *Equation* (5) *remains true even if* $s=0$ *for almost all* w *in the set* $\{w : y(0,\ w)\in I\}$. *Equivalently, equation* (6) *remains true with* $t_1=0$, *for each* $j_1\in I$. Corollary 1 to Theorem 3 is similarly extended.

This corollary is the most inclusive form of the strong Markovian property in the chain case.

9. The following complement is useful. Two families of sets are said to be independent in case any set from one family is independent of any set from the other family. A family of random variables and a family of sets are said to be independent in case the (smallest augmented) Borel field generated by the family of random variables is independent of the given family of sets. The Borel field $\Delta\mathscr{F}\{y_t,\ t\geqq 0\}=\Delta\mathscr{F}\{y_t,\ t>0\}$ may also be denoted by $\Delta\mathscr{F}\{x_t,\ t\geqq\alpha\}$, cf. the first paragraph of this paper.

THEOREM 6. *A necessary and sufficient condition that the two Borel fields* $\mathscr{F}_\alpha$ *and* $\Delta\mathscr{F}\{y_t,\ t\geqq 0\}$ *be independent with respect to* $\mathrm{P}(\circ\,|\,\Delta)$ *is that for each* $j\in I$ *there exists a function* $\rho_j(\circ)$ *on* $(0,\ \infty)$ *such that for every* $\Lambda\in\mathscr{F}$ *we have*

$$r_j(s,\ t\,|\,\Lambda)=\rho_j(t-s) \tag{25}$$

for a.a. $s\geqq 0$ (*in* $A(\Lambda\ ;\ \circ)$ *measure*) *and* $t>s$. *The function* $\rho_j(\circ)$ *is then identified to be* $r_j(\Delta\ ;\ \circ)/\mathrm{P}(\Delta)$.

PROOF. Suppose that (25) holds ; then we have by (20), if $0<t_1<$

$\cdots < t_N\,;\, j_1, \cdots, j_N \in I,$

$$\mathbf{P}\{\Lambda\,;\, y(t_v, w) = j_v,\ 1 \leqq v \leqq N \mid \Delta\} = \mathbf{P}(\Lambda \mid \Delta)\rho_{j_1}(t_1)R\ .$$

Taking the particular case $\Lambda = \Delta$ and comparing the two equations we obtain

$$\begin{aligned}\mathbf{P}\{\Lambda\,;\, y(t_v, w) = j_v,\ 1 \leqq v \leqq N \mid \Delta\} \\ = \mathbf{P}(\Lambda \mid \Delta)\mathbf{P}\{y(t_v, w) = j_v,\ 1 \leqq v \leqq N \mid \Delta\}\ .\end{aligned}$$

This implies the asserted independence ; the identification of ρ_j follows from (21).

Conversely, suppose that there is the asserted independence. Let $u \geqq 0$ and put $\Lambda_1 = \Lambda \cap \{w : \alpha(w) \leqq u\}$. We note that

$$A(\Lambda_1\,;\, s) = A(\Lambda\,;\, \min(s, u)).$$

Furthermore, by Corollary 1 to Theorem 3, we have

$$r_j(s, s + t \mid \Lambda) = \mathbf{P}\{y(t, w) = j \mid \Lambda\,;\, \alpha(w) = s\}\ , \quad t > 0, \text{ a.a. } s \geqq 0.$$

Using this interpretation[17] we see that

$$r_j(s, s + t \mid \Lambda) = r_j(s, s + t \mid \Lambda_1)\ , \qquad t > 0, \text{ a.a. } s \in [0, u].$$

It follows from (21) and these remarks that

$$\mathbf{P}\{\Lambda_1\,;\, y(t, w) = j\} = \int_{[0,u]} r_j(s, s + t \mid \Lambda) dA(\Lambda\,;\, s)\ .$$

On the other hand, the supposed independence implies that

$$\begin{aligned}\mathbf{P}\{\Lambda_1\,;\, y(t, w) = j\} &= \frac{1}{\mathbf{P}(\Delta)}\mathbf{P}(\Lambda_1)\mathbf{P}\{y(t, w) = j\} \\ &= \frac{1}{\mathbf{P}(\Delta)}\int_{[0,u]} r_j(\Delta\,;\, t) dA(\Lambda\,;\, s)\ .\end{aligned}$$

Comparing the last two displays we obtain, since u is arbitrary,

$$r_j(s, s + t \mid \Lambda) = \frac{1}{\mathbf{P}(\Delta)} r_j(\Delta\,;\, t) \qquad t > 0, \text{ a.a. } s \geqq 0.$$

By Fubini's theorem, the last equation holds if $0 \leqq s \notin Z$ and $0 \leqq t \notin Z_s$ where Z is of $A(\Lambda\,;\, \circ)$ measure zero and each Z_s is of Lebesgue measure zero. Replacing $s + t$ by t we obtain (25) if $0 \leqq s \notin Z$ and $s < t \notin s(+)Z_s$. But for a. a. s both members of (25) are continuous in $t \in (s, \infty)$ by Lemma 3 ; hence (25) is true if $0 \leqq s \notin Z$ and $t > s$.

COROLLARY. *The random variable α is independent of $\Delta\mathscr{F}\{y_t,\ t \geqq 0\}$*

[17] The reader is gravely warned against mistaking $r_j(s, s+t \mid \Lambda)$ to be $\mathbf{P}\{x(s+t, w) = j \mid \Lambda;\ \alpha(w) = s\}$ by any manner or means — the latter symbol is not defined.

if and only if, for each j, there exists a function $\rho_j(\circ)$ on $(0, \infty)$, such that $r_j(s, t|\Delta) = \rho_j(t-s)$ for a.a. $s \geqq 0$, $t > s$.

A particularly important case in applications is where α is such that for a certain $j \in I$ we have

$$r_j(\Delta\,;\,0) = \mathbf{P}\{y(0, w) = j\} = \mathbf{P}(\Delta)\,. \tag{26}$$

This is the case where $\underline{\lim}_{t \downarrow \alpha(w)} x(t, w) = j$ for almost all $w \in \Delta$. The content of Theorem 5, can then be stated as follows. If $\Lambda \in \mathscr{F}_\alpha$, $0 = t_1 < t_2 < \cdots < t_N$; $j_1 = j$ and $j_2, \cdots, j_N \in \overline{I}$, we have

$$\mathbf{P}\{\Lambda\,;\, y(t_v, w) = j_v,\ 1 \leqq v \leqq N\} = \mathbf{P}(\Lambda)\prod_{v=1}^{N-1} p_{j_v j_{v+1}}(t_{v+1} - t_v)\,.$$

This is the extension of (6) to cover $t_1 = 0$. There is thus "absolute independence" of the past $\mathscr{F}_\alpha$ and the future $\mathscr{F}\{y_t, t \geqq 0\}$ and the post-α process $\{y_t, t \geqq 0\}$ starts "from scratch" at the state j. In particular, $r_k(\Delta\,;\,t) = \mathbf{P}(\Delta)p_{jk}(t)$ for every $k \in \overline{I}$. Hence we have by Theorem 6,

$$\mathbf{P}\{x(t, w) = k\,|\,\alpha(w) = s\} = p_{jk}(t-s) \qquad \text{a.a. } s \geqq 0,\ t > s. \tag{27}$$

To illustrate the use of (27), let $i \neq j$ and define on the set $\{w: x(0, w) = i\}$

$$\alpha_{ij}(w) = \inf\{t \geqq 0\,;\, x(t, w) = j\}\ .$$

It is easily seen, using (ii), that α_{ij} is an optional random variable. Let

$$F_{ij}(t) = \mathbf{P}\{\alpha_{ij}(w) \leqq t\}\ , \qquad t \geqq 0.$$

We have then, using (27),

$$\begin{aligned}\mathbf{P}\{\alpha_{ij}(w) \leqq t\,;\, x(t, w) = k\} &= \int\{\mathbf{P}[x(t, w) = k\,|\,\alpha_{ij}(w)]\,\|\,\alpha_{ij}(w) \leqq t\} \\ &= \int_{[0,t]} p_{jk}(t-s)dF_{ij}(s)\ .\end{aligned}$$

Formula (28) is a very special case of a general formula which will not be discussed here. Such formulas have frequently been regarded as obvious; it is well to keep in mind what a (rigorous) proof involves.

Another case of Theorem 6, is the post-exit process discussed in detail in [2]. We recall that it leads to differentiability properties of the functions p_{ij} which were discovered only recently. The more general results obtained in the present paper should also have analytical repercussions, perhaps even in the semi-group theory of Markov matrices.

ADDENDA

(May 1958). A considerably simpler proof of the full result (Corollary to Theorem 5) has been found which will be sketched here. Observe

first that Theorem 4 can be proved without Theorem 3. We have, using the notations in (6) but with $t_1 \geqq 0$,

$$\Lambda \cap \{w : y(t_v, w) = j_v, 1 \leqq v \leqq N\} = \lim_n \sup \Lambda_n$$

up to a set of probability zero, where

$$\Lambda_n = \Lambda \cap \cup_{m=0}^{\infty} \left\{ w : \frac{m}{n} \leqq \alpha(w) < \frac{m+1}{n}; \; x\left(\frac{m+1}{n} + t_v, w\right) = j_v, 1 \leqq v \leqq N \right\}.$$

This follows from separability and the right lower semi-continuity of $x(\circ, w)$. By the method of Section 7, we obtain

$$\mathbf{P}(\Lambda_n) \geqq \int_0^\infty r_{j_1}\left(s, \frac{[ns+1]}{n} + t_1 \mid \Lambda\right) R\, dA(\Lambda ; s) .$$

It follows from Lemma 3 that

$$\mathbf{P}\{\Lambda ; y(t_v, w) = j_v, 1 \leqq v \leqq N\} \geqq \lim_{n \to \infty} \mathbf{P}(\Lambda_n) = r_{j_1}(\Lambda ; t_1) R .$$

Summing over all $j_1, \cdots, j_N$ and using Theorem 4 if $t_1 = 0$, we see that equality must hold above, proving (6).

Furthermore, it can be shown that the post-α process as it stands is separable with any denumerable set everywhere dense in T. This is important for applications. For the proof we need the continuity property of $r_j(s, \circ | \Lambda)$ given in Lemma 3. The latter property, though stated here as part of a lemma appears to be essential for certain applications of the strong Markov property. Details of the above statements and other improvements will be given in a forthcoming monograph on Markov chains.

Just before the galley proofs of this paper were received I received a copy of the Russian journal "Theory of probability and its applications", Vol. II, No. 2 (1957). Results overlapping mine were obtained by Yushkevič in his paper "On strong Markov processes" pp. 187–213. His earlier paper in Učenye Zapiski Moscow University, Vol. 9, is inaccessible here.

UNIVERSITY OF CHICAGO AND SYRACUSE UNIVERSITY

BIBLIOGRAPHY

0. R. M. BLUMENTHAL, *An extended Markov property*, Trans. Amer. Math. Soc., 85 (1957) 52–72.
1. K. L. CHUNG, Foundations of the theory of continuous parameter Markov chains, *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, 1956, 29–40.

2. ———, *Some new developments in Markov chains*, Trans. Amer. Math. Soc., 81 (1956), 195–210.
3. J. L. DOOB, *Topics in the theory of Markoff chains*, Trans. Amer. Math. Soc., 52 (1942), 37–64.
4. ———, *Markoff chains — denumerable case*, Trans. Amer. Math. Soc., 58 (1945), 455–473.
5. ———, Stochastic processes, New York, 1953.
6. E. DYNKIN and A. JUSHKEVICH, *Strong Markov processes* (in Russian with English summary), Теория Бероятностей иее Применения, 1 (1956), 149–155.
7. W. FELLER and H. P. MCKEAN, JR., *A diffusion equivalent to a countable Markov chain*, Proc. Nat. Acad. Sci., U. S. A., 42 (1956), 351–354.
8. E. B. HOBSON, Theory of Functions of a real variable, Vol. 1, 3rd Ed. Dover Publications, New York.
9. G. A. HUNT, *Some theorems concerning Brownian motion*, Trans. Amer. Math. Soc., 81 (1956), 294–319.
10. P. LÈVY, *Systèmes Markoviens et stationaires. Cas dénombrables*. Ann. Ècole Norm. Sup. 68 (1951), 327–381.
11. D. RAY, *Stationary Markov processes with continuous paths*, Trans. Amer. Math. Soc., 82 (1956), 452–493.
12. E. C. TITCHMARSH, Theory of Functions, Oxford, 1939.

Reprinted with permission from Proceedings of the International Congress of Mathematicians, 14–21 August 1958 (Cambridge University Press, 1960).

CONTINUOUS PARAMETER MARKOV CHAINS†

By KAI-LAI CHUNG

The name given in the title is an abbreviation of 'Markov processes with continuous time parameter, denumerable state space and stationary transition probabilities'. This theory is to the discrete parameter theory as functions of a real variable are to infinite sequences. New concepts and problems arise which have no counterpart in the latter theory. Owing to the sharply defined nature of the process, these problems are capable of precise and definitive solutions, and the methodology used well illustrates the general notions of stochastic processes. It is possible that the results obtained in this case will serve as a guide in the study of more general processes. The theory has contacts with that of martingales and of semi-groups which have been encouraging and may become flourishing. For lack of space the developments from the standpoint of semi-groups or systems of differential equations cannot be discussed here.

Terms and notation not explained below follow more or less standard usage such as in [11].

Let $(\Omega, \mathfrak{F}, \mathbf{P})$ be a probability triple where $(\mathfrak{F}, \mathbf{P})$ is complete; $\mathbf{T} = [0, \infty)$, $\mathbf{T}^0 = (0, \infty)$, $\mathfrak{B}$ the usual Borel field on $\mathbf{T}$. Let $\{x_t, t \in \mathbf{T}\}$ be a Markov chain with the minimal state space $\mathbf{I}$, the stationary transition matrix $\{(p_{ij})\}$ and an arbitrary fixed initial distribution. By the minimal state space we mean the smallest denumerable set (of real numbers) such that $\mathbf{P}\{x_t(\omega) \in \mathbf{I}\} = 1$ for every $t \in \mathbf{T}$. The transition matrix is characterized by the following properties: for every $i, j \in \mathbf{I}$, $s, t \in \mathbf{T}^0$:

$$p_{ij}(t) \geqslant 0, \quad \sum_{j \in \mathbf{I}} p_{ij}(t) = 1, \quad p_{ij}(s+t) = \sum_{k \in \mathbf{I}} p_{ik}(s) p_{kj}(t). \tag{1}$$

The last of these relations is the semi-group property. In order to have a separable and measurable Markov chain with the given $\mathbf{I}$ and $\{(p_{ij})\}$ it is sufficient (and essentially necessary) that

$$\lim_{t \downarrow 0} p_{ii}(t) = 1 \tag{2}$$

for every $i \in \mathbf{I}$. In this case we define $p_{ij}(0) = \delta_{ij}$. Each p_{ij} is then uniformly continuous in $\mathbf{T}$. We shall confine ourselves to this case. We may suppose, by going to a *standard modification*, that the process is separable

† This paper was prepared with the partial support of the Office of Scientific Research of the United States Air Force.

(relative to the class of closed sets) and measurable. The basic sample function properties, due largely to Doob[9] and Lévy[17], can then be deduced. We cannot detail these properties but, as a consequence of them, there is a specific *version* (or *realization*) which has desirable properties. To obtain this let us take $\mathbf{I}$ to be the set of positive integers and compactify it by adjoining one fictitious state ∞, in the usual manner. Let us call a function f on $\mathbf{T}$ *lower right semi-continuous* with respect to a denumerable set R dense in T if $f(t) = \underline{\lim}_{R \ni s \downarrow t} f(s)$ for every t. Then there is a version of the given Markov chain such that

(i) $x(\cdot,\cdot)$ is measurable $\mathfrak{B} \times \mathfrak{F}$, or Borel measurable;

(ii) each sample function $x(\cdot,\omega)$ is lower right semi-continuous with respect to any R.

Other properties which are valid for almost all sample functions may be further imposed; we need not elaborate them here. Clearly (i) implies measurability and (ii) implies well-separability, namely separability with respect to any denumerable set dense in $\mathbf{T}$. We mention that, despite (ii), it is possible that for all ω, the t-set $S_\infty(\omega)$ where $x(t,\omega) = \infty$ is everywhere dense in $\mathbf{T}$ (see [14]).

It has been known for some time that an important concept in the study of general Markov processes is the so-called strong Markov property (see [3, 12, 22]). It turns out that the version specified above has this property, which we now proceed to describe (in a slightly restricted form). Let $\mathfrak{F}_t$ be the Borel subfield of $\mathfrak{F}$ generated by $\{x_s, 0 \leqslant s \leqslant t\}$ and augmented by all sets of probability zero. A random variable α with domain of finiteness Δ_α is said to be *optional* (or 'stopping time' or 'independent of the future') if for every t we have

$$\{\omega : \alpha(\omega) < t\} \in \mathfrak{F}_t.$$

The collection of sets Λ in $\mathfrak{F}$ such that $\Lambda \cap \{\omega : \alpha(\omega) < t\} \in \mathfrak{F}_t$ is a Borel field $\mathfrak{F}_\alpha$ called the *pre-α field*. The process $\{\xi_t, t \in \mathbf{T}\}$ on the reduced triple $(\Delta_\alpha, \Delta_\alpha \mathfrak{F}, \mathbf{P}(\cdot \mid \Delta_\alpha))$ where

$$\xi(t,\omega) = x(\alpha(\omega)+t, \omega)$$

is called the *post-α process* and the augmented Borel field generated by this process the *post-α* field $\mathfrak{F}'_\alpha$. Observe that if α is optional then so is $\alpha + t$ for each $t > 0$. For the sake of brevity we shall suppose that $\Delta_\alpha = \Omega$ in the following. The following assertions, collectively referred to as the strong Markov property here, are true for every optional α.

(*a*) For each $t \in \mathbf{T}^0$, ξ_t is finite (i.e. $\xi_t \in \mathbf{I}$) with probability one.

(*b*) The post-α process is a Markov chain in $\mathbf{T}^0$ whose state space and transition matrix are restrictions of those of the given Markov chain.

(*c*) For each $t \in \mathbf{T}$, the pre-$(\alpha+t)$ and post-$(\alpha+t)$ fields are conditionally independent given ξ_t, wherever the latter is finite (hence in particular almost everywhere if $t \in \mathbf{T}^0$, by (*a*)). Thus if $\Lambda \in \mathfrak{F}_{\alpha+t}$ and $\Lambda' \in \mathfrak{F}'_{\alpha+t}$ then on the set $\{\omega: \xi_t(\omega) \in \mathbf{I}\}$ we have

$$\mathbf{P}\{\Lambda\Lambda' \mid \xi_t\} = \mathbf{P}\{\Lambda \mid \xi_t\}\,\mathbf{P}\{\Lambda' \mid \xi_t\}.$$

(*d*) The post-α process is well-separable and Borel measurable as it stands.

Furthermore, let us consider for each given $\Lambda \in \mathfrak{F}_\alpha$ and $s \leqslant t$, the conditional probability

$$\mathbf{P}\{x(t,\omega) = j \mid \Lambda;\, \alpha = s\} = r_j(s,t \mid \Lambda) \tag{3}$$

defined for almost all s according to the measure induced by α on $\mathfrak{B}$. The following additional assertions are true.

(*e*) For each $j \in \mathbf{I}$ and almost all s according to the α-measure: the function $r_j(s,\cdot \mid \Lambda)$ satisfies conditions analogous to (1) and is continuous in (s,∞); and we have

$$\mathbf{P}\{\xi(t,\omega) = x(\alpha(\omega)+t,\omega) = j \mid \Lambda;\, \alpha = s\} = r_j(s, s+t \mid \Lambda). \tag{4}$$

(*f*) The pre-α and post-α fields are absolutely independent if and only if $r_j(s,t \mid \Lambda)$ as a function of the pair (s,t) is a function of $t-s$, for each Λ and j.

Let us observe, comparing (3) and (4), that the assertion (*e*) is a non-trivial substitution property for the conditional probability.

A preliminary view of the above assertions, together with a justification of the name 'strong Markov property', may be obtained by considering the particular case α = constant. In this case the assertions (*a*)–(*d*) become the defining properties of $\{x_t, t \in \mathbf{I}\}$, while (*e*) reduces to the continuity of each p_{ij}. This simple observation implies the truth of (*a*)–(*e*) if α is denumerably-valued and shows that for a discrete parameter Markov chain the corresponding assertions hold almost trivially. Proofs of the above assertions except (*d*), in somewhat more precise terms, are given in [7]. Similar results which overlap these are given by Yushkevič[22]†; for another proof of (*a*), (*b*) and part of (*c*) see Austin[2].

The essence of the strong Markov property may be briefly stated as follows: The ordinary Markov property valid at a fixed time t remains valid at a variable time α chosen according to the evolution of the process but without prevision of the future. The classical illustration is that of a gambler who chooses his turn of playing according to a

† His assertion involving another random variable $\geqslant \alpha$ and measurable $\mathfrak{F}_\alpha$ follows easily from (*e*).

gambling system which he has devised without the aid of prescience. Similar concepts for a martingale have been developed by Doob[11].

Let us discuss some applications of the strong Markov property. It should be remarked that while its invocation is a basic step in each of the following cases further work is needed to establish the results to be mentioned.

(A) The simplest case of an optional random variable is the first entrance time into a given state. This satisfies (f) and its judicious use yields various 'decomposition formulas'. For example, let ${}_Hp_{ij}(t)$ denote the transition probability from i to j in time t 'under the taboo H' (namely, before entering any state in H), and ${}_{k,H}p_{ij}(t)$ the analogous probability where H is replaced by the union of H and k; ${}_HF_{ik}$ the first entrance time distribution from i to k under the taboo H. The intuitive meaning of the following formula is clear: if $k \notin H$, we have

$$ {}_Hp_{ij}(t) = {}_{k,H}p_{ij}(t) + \int_0^t {}_Hp_{kj}(t-s)\, d\, {}_HF_{ik}(s); $$

but its rigorous proof requires the strong Markov property, in particular (e). Specialization of H to one state leads to ratio limit theorems of the Doeblin type concerning $\int_0^t p_{jj}(s)\,ds \Big/ \int_0^t p_{ii}(s)\,ds$ as $t \to \infty$; see [8].

Next, let us recall that the state i is called *stable* or *instantaneous* according as $q_i = -p'_{ii}(0)$, which always exists, is finite or infinite. Let $i \neq j$ and consider in a recurrent class (see [8]) the successive returns to i via j (the intervention of j is necessary only if i is instantaneous). These return times partition the time axis $\mathbf{T}$ into independent blocks to which Doeblin's method of treating a functional of the Markov chain can be applied. In this way the classical limit theorems, like the laws of large numbers, the law of iterated logarithm and the central limit theorem, can be easily extended. For the discrete case see [4], where there are some errors in the proofs which can be corrected (see the last footnote in [8]).

Finally, Kolmogorov's example[16], in which there is exactly one instantaneous state, can be analysed probabilistically by use of certain entrances into this state and taboo probabilities. It can be shown as a consequence that the construction of sample functions of this process given by Kendall and Reuter[15] with semi-group methods is indeed the unique one. Namely, the version specified above having Kolmogorov's transition matrix must have the properties implied by the Kendall–Reuter construction.

(B) Let the process start at a stable state i and consider the first exit time from i (see [6]). This is optional and the condition in (f) is again satisfied. Denoting for $\Lambda = \Omega$ the corresponding $r_j(s, s+t)$ by $r_{ij}(t)$, which is continuous in t by (e), we have readily

$$p_{ij}(t) = \delta_{ij}\, e^{-q_i t} + q_i \int_0^t e^{-q_i(t-s)}\, r_{ij}(s)\, ds.$$

This integral representation implies the existence of a continuous derivative $p'_{ij}(t) = q_i[r_{ij}(t) - p_{ij}(t)]$ and various complements including an interpretation of Kolmogorov's first (backward) system of differential equations. (The second (forward) system can be dually treated and falls under (A) above.) This gives a probabilistic proof of a result which was first established by analytic means by Austin[1]. A proof similar to the one sketched here was announced by Yushkevič but has not yet appeared.

The independence of the pre-exit and post-exit fields implies the fundamental observation due to Lévy[17] that the lengths of the stable intervals are independently distributed; the separability of the post-exit process asserted in (d) then yields the negative exponential distributions for these lengths; see [5].

(C) Let the process start at an instantaneous state i and put

$$S_i(\omega) = \{t : x(t, \omega) = i\}, \quad \mu_i(t, \omega) = \mu[S_i(\omega) \cap (0, t)],$$

where μ is the Borel–Lebesgue measure on $\mathfrak{B}$. Then for each s, the random variable α_s defined by

$$\alpha_s(\omega) = \inf\{t : \mu_i(t, \omega) > s\}$$

is optional. In words, α_s is the first time when the total amount of time spent in the state i exceeds s. This idea, which is a partial analogue of the exit time from a stable state discussed under (B), is due to Lévy[17, 18, 19]. Lévy makes use of the more general device of counting time only on a selected set of states, thereby annihilating the remaining states together with the time spent in them. This idea remains to be fully exploited. As a simple example, if $i \neq j$, then the total time spent in i before entering j has the negative exponential distribution $1 - e^{-a_{ij} t}$, where

$$a_{ij} = \int_0^\infty {}_j p_{ii}(t)\, dt$$

in our previous 'taboo' notation.

(D) In this last application we touch upon a chapter of the theory of continuous parameter Markov chains which has yet to be written. It is to be observed that the strong Markov property fails on the set where $\xi_0 = \infty$. While the assertions (a)–(d) are always valid for $t \in \mathbf{T}^0$, our information is inadequate as the critical time α is approached from the

right if the sample function values approach ∞ there. This failure may be formally attributed to the crude one-point compactification we have adopted, which does not distinguish between the various modes of approaching ∞ that ought to correspond to distinct adjoined fictitious states rather than the one and only ∞. From this point of view the main task, called the 'boundary problem' by certain authors, is the proper compactification of the minimal state space **I** so as to restore the strong Markov property on the set where $\xi_0 \notin \mathbf{I}$ and to induce the appropriate boundary behavior (as in potential theory). Without loss of generality we may suppose that this set has probability one. For fixed $j \in \mathbf{I}$ and $t_0 \in \mathbf{T}^0$ let us consider the process $\{\eta_t, 0 \leqslant t \leqslant t_0\}$, where $\eta(t, \omega) = p_{\xi(t,\omega),j}(t_0 - t)$. Since $\{\xi_t\}$ is a Markov chain by (*b*) the new process $\{\eta_t\}$ is easily seen to be a martingale. Applying the martingale convergence theorem we see that $\lim_{t\downarrow 0} \eta(t, \omega)$ exists and is finite with probability one, and the limit has certain gratifying properties. The idea of considering this sort of martingale is due to Doob[9], and the present application to the post-α process will undoubtedly play a role in the compactification problem. For other formulations of the boundary problem see Feller[13], Reuter[21], and Ray[20].

The preceding discussion is centered around the strong Markov property as a convenient rallying point. Lest the impression should have been made that there was nothing else to be done I should like to conclude my discussion by mentioning some other problems not directly connected with the above.

A very natural circle of problems concerns the analytical properties (not to say characterization) of the elements of a transition matrix defined by (1) and (2). These may be regarded as problems in pure analysis. For example, it is still an open problem whether $p'_{ij}(t)$ exists if $t > 0$ and both i and j are instantaneous.† The solution of such a problem would be the more interesting if probabilistic significance is found. In this connection Jurkat[14] has observed that the differentiability results discussed under (B) hold even if the second condition in (1) is omitted, the condition (2) being assumed of course. The following even more primitive and probabilistically meaningful result is only a few weeks old: each p_{ij} is either identically zero or never zero. The original proof of this result, due to Austin, makes ingenious use of the strong Markov property. It is almost 'unfortunate' that a simplification has been found by myself which uses only the separability and measurability of an associated process. This is not the only example where a purely analytic

† (*Added in proof.*) D. Ornstein has now proved that for every i and j, $p'_{ij}(t)$ exists and is continuous for $t > 0$.

and simple-sounding proposition has so far been proved only by properly probabilistic methods.† It follows from this result, as observed by Austin, that if all states communicate then each of the Kolmogorov systems holds as soon as one of its equations holds for one value of the argument.

Another circle of problems is the approximation of the continuous parameter chain $\mathfrak{C} = \{x_t, t \in \mathbf{T}\}$ by its *discrete skeletons* $\mathfrak{C}_s = \{x_{ns}, n \in \mathbf{N}\}$ where $s > 0$ and $\mathbf{N}$ is the sequence of non-negative integers. In what sense and how well do the skeletons $\mathfrak{C}_s$ approximate $\mathfrak{C}$ as $s \downarrow 0$? This does not appear to be as simple a matter as might be expected. To cite a specific example: let m_{ij} denote the mean first entrance time (or return time if $i = j$) from i to j in $\mathfrak{C}$, and let $m_{ij}(s)$ denote the analogous quantity in $\mathfrak{C}_s$. The well-known theorem that $\lim_{t\to\infty} p_{ii}(t)$ exists (see [17]) implies that if i is stable then $m_{ii}(s) = q_i m_{ii}$ for every s. If i and j are distinct states in a positive (or strongly ergodic) class then it can be shown that $\lim_{s\downarrow 0} s m_{ij}(s) = m_{ij}$ by a rather devious method. But I do not know what the situation is with moments of higher order. We may also mention the open problem of characterizing a discrete parameter Markov chain which can be imbedded in a continuous parameter one, namely which is a skeleton of the latter.

Finally, let me mention an annoying kind of problem. Various models of Markov chains can be easily described by so-to-speak word-pictures but the rigorous verification that they are indeed Markovian is often laborious. The well-known construction by Doob[10] is an example. Other examples are given by Lévy[17] of which one (his example II.10.5) may be roughly described as follows. Consider first the infinite *descending escalator* such that from the state $i+1$ one necessarily goes into the state i while the mean sojourn times in all the states form a convergent series, the process terminating at the state 1. This is a Markov chain in $\mathbf{T}^0$ and one need only hitch it on to a new state at the beginning to obtain a Markov chain in $\mathbf{T}$. In fact, the resulting process is the second example given by Kolmogorov[16], which like the first one mentioned under (A) above has been analysed in detail by Kendall and Reuter[15]. Now modify this scheme by allowing, upon leaving each step, the alternative of either entering the next lower step or starting all over again from the (infinite) top of the escalator. By proper choice of the probabilities of the alternatives it is possible to jump to and return from infinity a nondenumerably infinite number of times. It seems 'intuitively obvious' that the resulting process is still Markovian, but if so why does it elude a simple proof?

† (*Added in proof.*) D. Ornstein has now found an analytical proof of this result.

517 CONTINUOUS PARAMETER MARKOV CHAINS

REFERENCES

[1] Austin, D. G. On the existence of derivatives of Markoff transition probability functions. *Proc. Nat. Acad. Sci., Wash.*, 41, 224–226 (1955).

[2] Austin, D. G. A new proof of the strong Markov theorem of Chung. *Proc. Nat. Acad. Sci., Wash.*, 44, 575–578 (1958).

[3] Blumenthal, R. M. An extended Markov property. *Trans. Amer. Math. Soc.* 85, 52–72 (1957).

[4] Chung, K. L. Contributions to the theory of Markov chains. II. *Trans. Amer. Math. Soc.* 76, 397–419 (1954).

[5] Chung, K. L. Foundations of the theory of continuous parameter Markov chains. *Proceedings Third Berkeley Symposium on Mathematical Statistics and Probability*, 2, 29–40 (1955).

[6] Chung, K. L. Some new developments in Markov chains. *Trans. Amer. Math. Soc.* 81, 195–210 (1956).

[7] Chung, K. L. On a basic property of Markov chains. *Ann. Math.* (2), 68, 126–149 (1958).

[8] Chung, K. L. Some aspects of continuous parameter Markov chains. *Publ. Inst. Statist. Univ. Paris*, 6, 271–287 (1957).

[9] Doob, J. L. Topics in the theory of Markoff chains. *Trans. Amer. Math. Soc.* 52, 37–64 (1942).

[10] Doob, J. L. Markoff chains—denumerable case. *Trans. Amer. Math. Soc.* 58, 455–473 (1945).

[11] Doob, J. L. *Stochastic Processes*. New York, 1953.

[12] Dynkin, E. and Yushkevič, A. Strong Markov processes. *Theory of Probability and Its Applications*, 1, 149–155 (1956).

[13] Feller, W. On boundaries and lateral conditions for the Kolmogorov differential equations. *Ann. Math.* (2), 65, 527–570 (1957).

[14] Feller, W. and McKean, H. A diffusion equivalent to a countable Markov chain. *Proc. Nat. Acad. Sci., Wash.*, 42, 351–354 (1956).

[15] Kendall, D. G. and Reuter, G. E. H. Some pathological Markov processes with a denumerable infinity of states and the associated semi-groups of operators on l. *Proceedings International Congress of Mathematicians, Amsterdam*, 3, 377–415 (1954).

[16] Kolmogorov, A. N. On some problems concerning the differentiability of the transition probabilities in a temporally homogeneous Markov process having a denumerable set of states. *Učenye Zapiski (Matem.) Moskov. Gos. Univ.* (4), 148, 53–59 (1951).

[17] Lévy, P. Systèmes markoviens et stationnaires; cas dénombrables. *Ann. Sci. Éc. Norm.* (3), 68, 327–381 (1951).

[18] Lévy, P. Compléments à l'étude des processus de Markoff. *Ann. Sci. Éc. Norm.* (3), 69, 26–212 (1952).

[19] Lévy, P. Processus markoviens et stationnaires du cinquième type (infinité dénombrable des états possibles, paramètre continu). *C.R. Acad. Sci., Paris*, 236, 1630–1632 (1953).

[20] Ray. D. (To appear.)

[21] Reuter, G. E. H. Denumerable Markov processes and the associated contraction semi-groups on l. *Acta Math.* 97, 1–46 (1957).

[22] Yushkevič, A. On strong Markov processes. *Theory of Probability and Its Applications*, 2, 187–213 (1957).

PRINTED IN GREAT BRITAIN

Journal of the Mathematical Society of Japan Vol. 11, No. 4, October, 1959

On the Lipschitz's condition for Brownian motion.

By K. L. CHUNG*, P. ERDÖS and T. SIRAO

(Received April 9, 1959)
(Revised May 25, 1959)

Let $X(t)$ $(0 \leqq t < \infty)$ be the Brownian motion process. Concerning the uniform continuity of $X(t)$, there exists P. Lévy's result. Before stating his result, let us define the concept of upper class and lower class with regard to the uniform continuity of $X(t)$ $(0 \leqq t \leqq 1)$.

If there exists a positive number ε such that $|t'-t| \leqq \varepsilon$ implies the relation

$$|f(t')-f(t)| \leqq g(|t'-t|), \tag{1}$$

where $g(t)$ is a non-negative, continuous, non-decreasing function defined in some finite interval $(0, T)$ and vanishing with t, then we say that $f(t)$ satisfies Lipschitz's condition relative to $g(t)$. Putting $\varphi(t) = \psi\left(\frac{1}{t}\right)\sqrt{t}$, if $X(t)$ $(0 \leqq t \leqq 1)$ satisfies Lipschitz's condition relative to $\varphi(t)$ with probability 1 we say that $\psi(t)$ belongs to the upper class. If $X(t)$ $(0 \leqq t \leqq 1)$ does not satisfy Lipschitz's condition relative to $\varphi(t)$ with probability 1 we say that $\psi(t)$ belongs to the lower class. P. Lévy [1] proved that the function

$$\psi(t) = c(2 \log t)^{\frac{1}{2}}$$

belongs to the upper class for $c > 1$ and belongs to the lower class for $c < 1$. Following his method, T. Sirao [2] improved the result as follows: The function

$$\psi(t) = (2 \log t + c \log \log t)^{\frac{1}{2}}$$

belongs to the upper class for $c > 5$ and belongs to the lower class for $c < -1$. In this paper we shall prove the following theorems.

THEOREM 1. *A non-negative, continuous and monotone non-decreasing function $\psi(t)$ belongs to the upper or lower class according as the integral*

$$\int^{\infty} \psi^3(t) e^{-\frac{1}{2}\psi^2(t)} dt \tag{2}$$

is convergent or divergent.

---

* Research supported in part by the United States Air Force through the Office of Scientific Research under contract AF 49 (638)-265.

THEOREM 2. *The function* $\psi(t)$ *defined by*

$$\psi(t)=(2\log t+5\log_{(2)}t+2\log_{(3)}t+\cdots+2\log_{(n-1)}t+c\log_{(n)}t)^{\frac{1}{2}}, \tag{3}$$

where $\log_{(n)}t$ *denotes the n-times iterated logarithm, belongs to the upper class for* $c>2$ *and to the lower class for* $c\leqq 2$:

These theorems were quoted by P. Lévy [3] without proof. They give a difinitive solution to the problem of uniform continuity of Brownian motion $X(t)$ and are comparable to A. Kolmogorov's criterion in the theory of iterated logarithm for $X(t)$ at time point ∞.

Theorem 2 is a simple corollary of Theorem 1. Hence we prove only Theorem 1.

LEMMA 1. *Without loss of generality, we may assume that*

$$(2\log t-10\log\log t)^{\frac{1}{2}}\leqq\psi(t)\leqq(2\log t+10\log\log t)^{\frac{1}{2}}. \tag{4}$$

PROOF. We show that if Theorem 1 holds under the assumption (4), then it holds without (4). Let us denotes the first member in (4) by $\psi_1(t)$ and the last member in (4) by $\psi_2(t)$.

Define $\hat{\psi}(t)$ as follows:

$$\hat{\psi}(t)=\min(\max(\psi(t),\psi_1(t)),\psi_2(t)). \tag{5}$$

Then the convergence of the integral (2) for $\psi(t)$ implies the same for $\hat{\psi}(t)$. In fact, let us assume the convergence of (2) for $\psi(t)$. If the set of t on which $\psi(t)$ is less than $\psi_1(t)$ is not bounded, there exists an increasing sequence $\{t_n\}$ such that $\psi(t_n)\leqq\psi_1(t_n)$ and t_n tends to infinity with n. Since $\psi(t)$ is a non-negative and non-decreasing function, we have

$$\begin{aligned}\int_{t_1}^{\infty}\psi^3(t)e^{-\frac{1}{2}\psi^2(t)}dt &\geqq\int_{t_1}^{t_n}\psi^3(t)e^{-\frac{1}{2}\psi^2(t)}dt\\ &\geqq c\psi^3(t_n)e^{-\frac{1}{2}\psi^2(t_n)}t_n\\ &\geqq c(\log t_n)^{\frac{13}{2}}\end{aligned}$$

where c is a positive constant. Since $\log t_n$ tends to infinity with n, the integral for $\psi(t)$ is divergent. This contradicts our assumption and therefore $\psi_1(t)$ must be smaller than $\psi(t)$ for large t. On the other hand the integral for $\psi_2(t)$ is convergent. These facts prove our assertion. Now we assume that the integral for $\psi(t)$ is convergent and Theorem 1 valid under the condition (4). Then the integral for $\hat{\psi}(t)$ is convergent and therefore $\hat{\psi}(t)$ belongs to the upper class. But by what has just been shown $\hat{\psi}(t)\leqq\psi(t)$ for large t. So we have $\hat{\varphi}(h)\leqq\varphi(h)$ for small h where $\hat{\varphi}(t)$ is defined by $\hat{\psi}(t)$ as $\varphi(t)$ is by $\psi(t)$ and therefore $\psi(t)$ belongs to the upper class. Thus Lemma 1 is proved in the convergent case.

Secondly let us assume that the integral for $\psi(t)$ is divergent. If the set of t on which $\psi(t)$ is less than $\psi_1(t)$ is bounded, then it follows that $\hat{\psi}(t)$ is less than $\psi(t)$ for large t and accordingly the integral for $\hat{\psi}(t)$ must be divergent. On the contrary, if there exists an increasing sequence $\{t_n\}$ having the property

$$(6) \qquad \psi(t_n) < \psi_1(t_n), \quad t_n \to \infty \quad \text{as } n \to \infty,$$

then we have

$$(7) \qquad \hat{\psi}(t_n) = \psi_1(t_n).$$

By the monotony of $\hat{\psi}(t)$, we have

$$(8) \qquad \begin{aligned} \int_{t_1}^{t_n} \hat{\psi}^3(t) e^{-\frac{1}{2}\hat{\psi}^2(t)} dt &\geqq \psi^3(t_n) e^{-\frac{1}{2}\hat{\psi}^2(t_n)}(t_n - t_1) \\ &= \psi_1{}^3(t_n) e^{-\frac{1}{2}\psi_1^2(t_n)}(t_n - t_1). \end{aligned}$$

Since the last term in (8) tends to infinity with n, the integral for $\hat{\psi}(t)$ is divergent in our case. Now, by the result in [2], $\psi_2(t)$ belongs to the upper class and therefore, for almost all sample point ω, there exists ε such that

$$(9) \qquad |X(t', \omega) - X(t, \omega)| < \varphi_2(|t'-t|) \qquad \text{for} \quad |t'-t| < \varepsilon,$$

where $\varphi_2(t)$ is defined by $\psi_2(t)$ in the same way as $\varphi(t)$ is by $\psi(t)$. On the other hand, since by assumption $\psi_2(t)$ belongs to the lower class, for almost all ω we can choose a sequence $\{(t_n, t_n')\}$ having the following properties

$$(10) \qquad \begin{aligned} &|X(t_n') - X(t_n)| > \hat{\varphi}(|t_n' - t_n|), \\ &|t_n' - t_n| \to 0 \qquad \text{as} \quad n \to \infty. \end{aligned}$$

From (9) and (10), we have

$$(11) \qquad \hat{\varphi}(|t_n' - t_n|) < \varphi_2(|t_n' - t_n|).$$

(11) shows that $\varphi(t)$ is at last equal to $\varphi(t)$ at $t = |t_n' - t_n|$. This fact and (10) show that $\psi(t)$ belongs to the lower class. Q. E. D.

We now proceed to prove Theorem 1.

1) Proof of the convergent case.

First of all we remark that it suffices to prove, for almost all ω, the existence of a positive ε' such that

$$X(t', \omega) - X(t, \omega) \leqq \varphi(|t'-t|) \qquad \text{for} \quad |t'-t| < \varepsilon'.$$

In fact, let us assume that this assertion holds. Then it follows from the symmetry of Brownian motion that the probability of the existence of a positive ε'' satisfying the inequality

$$-\varphi(|t'-t|) \leqq X(t', \omega) - X(t, \omega) \qquad \text{for} \quad |t'-t| < \varepsilon''$$

is equal to 1. Taking ε for the minimum of ε' and ε'', we have Theorem 1. Therefore we may consider the difference $X(t')-X(t)$ instead of its absolute value.

For each triple (p, k, l), let $E^p_{k,l}$ be the event

$$X\left(\frac{k+l}{2^p}\right)-X\left(\frac{k}{2^p}\right)\geqq\varphi\left(\frac{l}{2^p}\right), \qquad \begin{aligned}&k=0,1,2,\cdots,2^p,\\ &l=1,2,\cdots,p.\end{aligned} \tag{12}$$

A simple computation shows that

$$P(E^p_{k,l})\sim e^{-\frac{1}{2}\psi^2\left(\frac{2^p}{l}\right)}\Big/(2\pi)^{\frac{1}{2}}\psi\left(\frac{2^p}{l}\right)$$

for large p. Summing up $P(E^p_{k,l})$ for $p=1,2,\cdots$, $k=1,2,\cdots,2^p$, $l=\left[\frac{p}{3}\right]$, $\left[\frac{p}{3}\right]+1,\cdots,p$, we have

$$\begin{aligned}\sum_{p=1}^{\infty}\sum_{k=1}^{2^p}\sum_{l=\left[\frac{p}{3}\right]}^{p}P(E^p_{k,l})&=O(1)\sum_{p=1}^{\infty}\sum_{k=1}^{2^p}\sum_{l=\left[\frac{p}{3}\right]}^{p}e^{-\frac{1}{2}\psi^2\left(\frac{2^p}{l}\right)}\Big/\psi\left(\frac{2^p}{l}\right)\\ &=O(1)\sum_{p=1}^{\infty}\sum_{k=1}^{2^p}\frac{p}{\psi\left(\frac{2^p}{p}\right)}e^{-\frac{1}{2}\psi^2\left(\frac{2^p}{p}\right)}.\end{aligned}$$

Applying Lemma 1, we obtain

$$\begin{aligned}\sum_{p=1}^{\infty}\sum_{k=1}^{2^p}\sum_{l=\left[\frac{p}{3}\right]}^{p}P(E^p_{k,l})&=O(1)\sum_{p=1}^{\infty}\frac{2^p}{p}\psi^3\left(\frac{2^p}{p}\right)e^{-\frac{1}{2}\psi^2\left(\frac{2^p}{p}\right)}\\ &=O(1)\int^{\infty}\psi^3(t)e^{-\frac{1}{2}\psi^2(t)}dt<+\infty.\end{aligned} \tag{13}$$

Next, for each triple (p, k, l), let $F^p_{k,l}$ be the event

$$\max_{0\leqq t,s\leqq\frac{1}{2^p}}\left\{X\left(\frac{k+l}{2^p}+t\right)-X\left(\frac{k}{2^p}-s\right)\right\}\geqq\sqrt{\frac{l}{2^p}}\,\psi\left(\frac{2^p}{l+2}\right), \tag{14}$$

$$k=0,1,2,\cdots,2^p,$$
$$l=1,2,\cdots,p.$$

For convenience' sake, we consider the $F^p_{k,l}$ only such that the time parameters t of $X(t)$ which appear in the above definition are positive and less than 1. It is well known that

$$P(\max_{0\leqq s\leqq t}X(s)>a)\leqq 2P(X(t)>a),$$

where a is any real number. Since the Brownian motion is an additive process, we have

$$(15)\quad \begin{aligned} P(F_{k,l}^p) &\leq P\left\{\max_{0\leq t\leq\frac{1}{2^p}}\left(X\left(\frac{k+l}{2^p}+t\right)-X\left(\frac{k+l}{2^p}\right)\right)+\left(X\left(\frac{k+l}{2^p}\right)-X\left(\frac{k}{2^p}\right)\right)\right.\\ &\qquad\left.+\max_{0\leq s\leq\frac{1}{2^p}}\left(X\left(\frac{k}{2^p}\right)-X\left(\frac{k}{2^p}-s\right)\right)\geq\sqrt{\frac{l}{2^p}}\,\psi\left(\frac{2^p}{l+2}\right)\right\}\\ &\leq 4P\left\{X\left(\frac{k+l+1}{2^p}\right)-X\left(\frac{k+l}{2^p}\right)+X\left(\frac{k+l}{2^p}\right)-X\left(\frac{k}{2^p}\right)\right.\\ &\qquad\left.+X\left(\frac{k}{2^p}\right)-X\left(\frac{k-1}{2^p}\right)\geq\sqrt{\frac{l}{2^p}}\,\psi\left(\frac{2^p}{l+2}\right)\right\}\\ &=4P\left(X\left(\frac{k+l+1}{2^p}\right)-X\left(\frac{k-1}{2^p}\right)\geq\sqrt{\frac{l}{2^p}}\,\psi\left(\frac{2^p}{l+2}\right)\right) \end{aligned}$$

By Lemma 1 we have, for large p and l,

$$\begin{aligned} P(F_{k,l}^p) &\leq \frac{4}{(2\pi)^{\frac{1}{2}}\psi\left(\frac{2^p}{l+2}\right)}\, e^{-\frac{l}{2(l+2)}\psi^2\left(\frac{2^p}{l+2}\right)}\\ &\sim 4P(E_{k,l}^p)\, e^{\frac{1}{l+2}\psi^2\left(\frac{2^p}{l+2}\right)}. \end{aligned}$$

Therefore, if l is an integer existing between $\left[\frac{p}{3}\right]$ and p, there exists a positive constant c such that

$$(16)\qquad P(F_{k,l}^p)\leq cP(E_{k,l}^p).$$

Combining (13) and (16), we obtain

$$(17)\qquad \sum_{p=1}^{\infty}\sum_{k=1}^{2^p}\sum_{l=\left[\frac{p}{3}\right]}^{p} P(F_{k,l}^p)<+\infty.$$

According to Borel-Cantelli's lemma in the convergent case, (17) shows that the events $F_{k,l}^p$ appearing in (17) occur "*only finitely many times*" with probability 1. Or, in other words, there exists a positive ε with probability 1 such that if $\frac{p}{2^{p+1}}$ is smaller than ε, $F_{k,l}^p$ does not occur for any pair (k,l) appearing in the summation of (17).

Now, for any pair of (t,t') satisfing the condition $|t'-t|<\varepsilon$, we choose p as follows:

$$(18)\qquad \frac{p+1}{2^{p+1}}<|t'-t|\leq\frac{p}{2^p}<2\varepsilon.$$

If we define k and l by the following inequalities

$$(19)\qquad \frac{k-1}{2^p}<\min(t,t')\leq\frac{k}{2^p}<\frac{k+l}{2^p}\leq\max(t,t')<\frac{k+l+1}{2^p},$$

it follows that $\left[\frac{p}{3}\right] < l \leqq p$ and therefore we obtain

$$
\begin{aligned}
X(t')-X(t) &\leqq \max_{0\leqq t,s\leqq\frac{1}{2^p}}\left(X\left(\frac{k+l}{2^p}+t\right)-X\left(\frac{k}{2^p}-s\right)\right)\\
&\leqq\left(\frac{l}{2^p}\right)^{\frac{1}{2}}\psi\left(\frac{2^p}{l+2}\right)\\
&\leqq\varphi(|t'-t|)
\end{aligned}
$$

with probability 1.

Thus Theorem 1 is proved in the convergent case.

2) Proof of the divergent case.

Let $E^p_{k,l}$ be the event defined by (12). By the monotony of $\psi(t)$ and Lemma 1, we have

$$
\begin{aligned}
\sum_{p=1}^{\infty}\sum_{k=1}^{2^p}\sum_{l=\left[\frac{p}{3}\right]+1}^{p}P(E^p_{k,l}) &= O(1)\sum_{p=1}^{\infty}\sum_{k=1}^{2^p}\sum_{l=\left[\frac{p}{3}\right]+1}^{p}e^{-\frac{1}{2}\psi^2\left(\frac{2^p}{l}\right)}\Big/\psi\left(\frac{2^p}{l}\right)\\
&= O(1)\sum_{p=1}^{\infty}\frac{2^p}{p}\psi^3\left(\frac{2^{p+1}}{p}\right)e^{-\frac{1}{2}\psi^2\left(\frac{2^{p+1}}{p}\right)} \qquad (20)\\
&= O(1)\int^{\infty}\psi^3(t)\,e^{-\frac{1}{2}\psi^2(t)}\,dt=+\infty\,.
\end{aligned}
$$

It is sufficient to show that $E^p_{k,l}$ occur "*infinitely often*" with probability 1. For this purpose, we use the following Lemma given in [4].

LEMMA 2. *Let $\{E_k\}$ be a sequence of events satisfying the following conditions.*

(i) $$\sum_{k=1}^{\infty}P(E_k)=+\infty\,.$$

(ii) *For every pair of positive integers h, n with $n\geqq h$, there exist $c(h)$ and $H(n,h) > n$ such that for every $m\geqq H(n,h)$ we have*

$$P(E_m/E_k',\cdots,E_n')>c(h)P(E_m)\,,$$

where $P(F/E)$ denotes the conditional probability of F on the hypothesis E and E' denotes the complement of E.

(iii) *There exist two absolute constants c_1 and c_2 with the following property: to each E_j there corresponds a set of events $E_{j_1},\cdots,E_{j_s}$ belonging to $\{E_k\}$ such that*

(a) $$\sum_{i=1}^{s}P(E_jE_{j_i})<c_1P(E_{j_i})$$

and if $k>j$ but E_k is not among the E_{j_i} $(1\leqq i\leqq s)$ then

(b) $$P(E_jE_k)<c_2P(E_j)P(E_k)\,.$$

Then the probability that the events E_k occur "infinitely often" is equal to one.

We rearrange $E_{k,l}^p$ and denotes it by E_m so that we may apply Lemma 2 in our case. The rule of ordering is given by the following. If $E_n = E_{k,l}^p$, $E_m = E_{k',l'}^{p'}$, then $n < m$ if and only if one of the following three conditions holds:

(α) $p < p'$,

(β) $p = p'$ and $l > l'$,

(γ) $p = p'$, $l = l'$ and $k < k'$.

Now we prove that the sequence $\{E_n\}$ satisfies the conditions of Lemma 2. (i) is a consequence of (20). For (ii), we use the characteristic property of Gaussian distribution. Let $E_m = E_{k,l}^p$ and put $U_m = X\left(\frac{k+l}{2^q}\right) - X\left(\frac{k}{2^p}\right)$. For every pair (h, n) with $n \geqq h$, if we define $U_h, U_{h+1}, \cdots, U_n$ similarly then

$$(21)\qquad \begin{aligned} &E(U_i) = 0 \quad (i = h, h+1, \cdots, n), \qquad E(U_m) = 0, \\ &E(U_i U_m) \leqq \frac{l}{2^p} \quad (i = h, h+1, \cdots, n), \end{aligned}$$

where $E(U)$ denotes the expectation of U. Since $\frac{l}{2^p}$ tends to zero as p increases, (21) shows that for each i $(h \leqq i \leqq n)$ the correlation coefficient of U_i and U_m tends to zero as m increases. In other words, U_m is asymptotically independent of the joint variable $(U_h, U_{h+1}, \cdots, U_n)$. Therefore we have

$$(22)\qquad \lim_{m\to\infty} \frac{P(E_m / E_h', \cdots, E_n')}{P(E_m)} = \lim_{m\to\infty} \frac{P(E_h', E_{h+1}', \cdots, E_n' / E_m)}{P(E_h', E_{h+1}', \cdots, E_n')} = 1.$$

This shows that (ii) holds in our case. For the justification of (iii), we need some lemmas.

LEMMA 3. *Let U and V be two random variables whose joint distribution is Gaussian and each of them has a standard Gaussian distribution. Let the correlation coefficient of U and V be ρ, then there exists a positive constant c_1 such that*

$$(23)\qquad P(U > a, V > b) \leqq c_1 P(U > a) P(V > b) \quad \text{for } \rho < \frac{1}{ab}.$$

PROOF. If ρ is negative or if a or b is small, (23) holds trivially. Therefore it is sufficient to prove Lemma 3 for sufficiently large a, b and positive ρ. Without loss of generality, we may assume $a \leqq b$. Then we have

$$P(U > a, V > b) = \frac{1}{2\pi(1-\rho^2)^{\frac{1}{2}}} \int_b^\infty \int_a^\infty e^{-\frac{(x^2 - 2\rho xy + y^2)}{2(1-\rho^2)}} dx dy$$

$$
\begin{aligned}
&= \frac{1}{2\pi(1-\rho^2)^{\frac{1}{2}}} \int_b^{2b}\int_a^{2b} e^{-\frac{(x-\rho y)^2}{2(1-\rho^2)}} e^{-\frac{y^2}{2}} dxdy \\
(24) \qquad &+ \frac{1}{2\pi(1-\rho^2)^{\frac{1}{2}}} \int_b^{\infty}\int_{2b}^{\infty} e^{-\frac{(x-\rho y)^2}{2(1-\rho^2)}} e^{-\frac{y^2}{2}} dxdy \\
&+ \frac{1}{2\pi(1-\rho^2)^{\frac{1}{2}}} \int_{2b}^{\infty}\int_a^{2b} e^{-\frac{(x-\rho y)^2}{2(1-\rho^2)}} e^{-\frac{y^2}{2}} dxdy .
\end{aligned}
$$

The first term on the right side is estimated as follows:

$$
\begin{aligned}
&\frac{1}{2\pi(1-\rho^2)^{\frac{1}{2}}} \int_b^{2b}\int_a^{2b} e^{-\frac{(x-\rho y)^2}{2(1-\rho^2)}} e^{-\frac{y^2}{2}} dxdy \\
(25) \qquad &\leqq \frac{1}{2\pi(1-\rho^2)^{\frac{1}{2}}} \int_b^{2b}\int_a^{2b} e^{-\frac{(x-2/a)^2}{2(1-\rho^2)}} e^{-\frac{y^2}{2}} dxdy \\
&\leqq P\{U > (a-2/a)/(1-\rho^2)^{\frac{1}{2}}\} P(V > b) \\
&= O(1) P(U > a) P(V > b) .
\end{aligned}
$$

On the other hand, for sufficiently large a, the second and third term on the right side of (24) are trivially smaller than the right side of (23) replaced c_1 by 1. These estimates assure the validity of Lemma 3. Q. E. D.

Lemma 4. *Let U and V be random variables as in Lemma 3. If the correlation coefficient of U and V is less than $1/2^{\frac{1}{2}}$ and $0 < a < b$ then there exist two positive constants c_2 and δ_2 satisfying the following inequality*

$$
(26) \qquad P(U > a,\ V > b) \leqq c_2 e^{-\delta_2 b^2} P(U > a) .
$$

Proof. Let ε be a positive constant which is less than 1 and let ρ be the correlation coefficient of U and V. It suffices to prove Lemma 4 for sufficiently large a and positive ρ. Then we have

$$
\begin{aligned}
P(U > a,\ V > b) &= \frac{1}{2\pi(1-\rho^2)^{\frac{1}{2}}} \int_a^{\infty}\int_b^{\infty} e^{-\frac{(x^2-2\rho xy+y^2)}{2(1-\rho^2)}} dxdy \\
(27) \qquad &= \frac{1}{2\pi(1-\rho^2)^{\frac{1}{2}}} \int_a^{(1+\varepsilon)^{\frac{1}{2}} b}\int_b^{\infty} e^{-\frac{(x-\rho y)^2}{2(1-\rho^2)}} e^{-\frac{y^2}{2}} dxdy \\
&+ \frac{1}{2\pi(1-\rho^2)^{\frac{1}{2}}} \int_{(1+\varepsilon)^{\frac{1}{2}} b}^{\infty}\int_b^{\infty} e^{-\frac{(x-\rho y)^2}{2(1-\rho^2)}} e^{-\frac{y^2}{2}} dxdy
\end{aligned}
$$

$$\leqq \frac{1}{2\pi(1-\rho^2)^{\frac{1}{2}}}\int_a^{(1+\varepsilon)^{\frac{1}{2}}b}\int_b^{\infty} e^{-\frac{(x-(1+\varepsilon/2)^{\frac{1}{2}})b}{2(1-\rho^2)}} e^{-\frac{y^2}{2}}dxdy$$

$$+\frac{1}{(2\pi)^{\frac{1}{2}}}\int_{(1+\varepsilon)^{\frac{1}{2}}b}^{\infty} e^{-\frac{y^2}{2}}dy$$

$$=O(1)\{e^{-(1-(1+\varepsilon/2)^{\frac{1}{2}})^2b^2/2}+e^{-\varepsilon/2b^2}\}P(U>a).$$

If we take the minimum of $(1-(1+\varepsilon/2)^{\frac{1}{2}})^2/2$ and $\varepsilon/2$ for δ_2 then Lemma 4 follows from (27) immediately. Q. E. D.

LEMMA 5. *Let U and V be random variables as in Lemma 3. Denoting the correlation coefficient of U and V by ρ, there exist two positive constants c_3 and δ_3 such that*

(28) $$P(U>a,\ V>a)\leqq c_3e^{-\delta_3(1-\rho^2)a^2}P(U>a) \qquad \text{for } a>0.$$

PROOF. By the definition of Gaussian distribution, we have

$$P(U>a,\ V>a)=\frac{1}{2\pi(1-\rho^2)^{\frac{1}{2}}}\int_a^{\infty}\int_a^{\infty} e^{-\frac{(x^2-2\rho xy+y^2)}{2(1-\rho^2)}}dydx.$$

Rotating the axes by $\pi/4$, we obtain

$$P(U>a,\ V>a)=\frac{1}{2\pi(1-\rho^2)^{\frac{1}{2}}}\int_{2^{\frac{1}{2}}a}^{\infty}\int_{-(x-2^{\frac{1}{2}}a)}^{(x+2^{\frac{1}{2}}a)} e^{-\frac{(1-\rho)x^2+(1+\rho)y^2}{2(1-\rho^2)}}dydx$$

$$\leqq \frac{1}{(2\pi)^{\frac{1}{2}}}\int_{(2/1+\rho)^{\frac{1}{2}}a}^{\infty} e^{-\frac{x^2}{2}}dx$$

$$=O(1)\,e^{-\frac{1-\rho^2}{2(1+\rho)^2}a^2}\cdot\frac{1}{a}\,e^{-\frac{a^2}{2}}$$

$$=O(1)\,e^{-\frac{1-\rho^2}{2(1+\rho)^2}}P(U>a).$$

If we take $1/8$ for δ_3, Lemma 5 follows from (31). Q. E. D.

Now we prove that the condition (iii) of Lemma 2 is satisfied by our sequence $\{E_n\}$. For given E_j, recalling that E_j has another expression $E_{k,l}^p$, we choose a sequence $\{E_{j_i};\, i=1,2,\cdots,s\}$ of events with the properties that $j_i>j$, the corresponding superscript p' is less than $(p+5\log p)$ and E_{j_i} is not independent of E_j. If E_m is independent of E_j then (b) of (ii) holds trivially for $c_2=1$. On the other hand, if E_m is not independent of E_j, we use Lemma 3. Let $E_j=E_{k,l}^p$ and $E_m=E_{k',l'}^{p'}$. If m is not one of the j_i's then it follows from the definition of $\{E_{j_i}\}$ that $(p+5\log p)<p'$. Considering only the case of $l>\frac{p}{2}$, we have by Lemma 1 and for large p,

$$
(30)\qquad E\left\{\frac{\left(X\left(\frac{k+l}{2^p}\right)-X\left(\frac{k}{2^p}\right)\right)}{\left(\frac{l}{2^p}\right)^{\frac{1}{2}}}\cdot\frac{\left(X\left(\frac{k'+l'}{2^{p'}}\right)-X\left(\frac{k'}{2^{p'}}\right)\right)}{\left(\frac{l'}{2^{p'}}\right)^{\frac{1}{2}}}\right\}\leqq\left(\frac{p'}{2^{p'}}\right)^{\frac{1}{2}}\left(\frac{2^{p+1}}{p}\right)^{\frac{1}{2}}
\leqq\frac{1}{\psi\left(\frac{2^p}{p}\right)\psi\left(\frac{2^{p'}}{p'}\right)}.
$$

Since the joint distribution of the two random variables appearing in (30) is a Gaussian distribution in 2-dimension's, we may use Lemma 3. Thus there exists a positive constant c such that

$$
(31)\qquad P(E_jE_m)=P\left\{X\left(\frac{k+l}{2^p}\right)-X\left(\frac{k}{2^p}\right)>\varphi\left(\frac{l}{2^p}\right),\ X\left(\frac{k'+l'}{2^{p'}}\right)-X\left(\frac{k'}{2^{p'}}\right)>\varphi\left(\frac{l'}{2^{p'}}\right)\right\}
\leqq cP(E_j)P(E_m).
$$

If we take the maximum of c and 1 for c_2 in (b) of (iii) then (b) holds.

In order to verify (a) of (iii), we use the other expressions of the E_j's. Let us denote E_j by $E^p_{k,l}$ and each one of E_{j_i} by E^{p}_{k,l_i}. Dividing the sum of $P(E_jE_{j_i})$ according to the magnitude of the correlation coefficient of $\left(X\left(\frac{k+l}{2^p}\right)-X\left(\frac{k}{2^p}\right)\right)$ and $\left(X\left(\frac{k'+l'}{2^{p'}}\right)-X\left(\frac{k'}{2^{p'}}\right)\right)$ we have

$$
(32)\qquad \sum_{i=1}^{s}P(E_jE_{j_i})=\Sigma'P(E_jE_{j_i})+\Sigma''P(E_jE_{j_i}),
$$

where Σ' denotes the summation over i's such that the correlation coefficient of the corresponding random variables is larger than $\frac{1}{\sqrt{2}}$ and Σ'' denotes the summation of the remainder. Since the correlation is at most

$$
\min\left(\frac{l}{2^p},\frac{l'}{2^{p'}}\right)\left(\frac{ll'}{2^{p+p'}}\right)^{-1/2}
$$

and since $l'2^{-p'}\leqq l2^{-p}$ by the limitation on the ranges of l and l', we see that the largest superscript of E_{j_i}'s appearing in Σ' is at most $p+2$. Moreover, without loss of generality, we may assume in the computation of $P(E_jE_{j_i})$ that $\frac{k}{2^p}\leqq\frac{k'}{2^{p'}}$. If $\frac{k+l}{2^p}\leqq\frac{k'+l'}{2^{p'}}$, we have

$$
(33)\qquad P(E_jE_{j_i})=P\left(\frac{X\left(\frac{k+l}{2^p}\right)-X\left(\frac{k}{2^p}\right)}{\left(\frac{l}{2^p}\right)^{\frac{1}{2}}}>\psi\left(\frac{2^p}{l}\right),\ \frac{X\left(\frac{k'+l'}{2^{p'}}\right)-X\left(\frac{k'}{2^{p'}}\right)}{\left(\frac{l'}{2^{p'}}\right)^{\frac{1}{2}}}>\psi\left(\frac{2^{p'}}{l'}\right)\right)
\leqq P\left(\frac{X\left(\frac{k+l}{2^p}\right)-X\left(\frac{k}{2^p}\right)}{\left(\frac{l}{2^p}\right)^{\frac{1}{2}}}>\psi\left(\frac{2^p}{l}\right),\ \frac{X\left(\frac{k+l}{2^p}\right)-X\left(\frac{k'}{2^{p'}}\right)}{\left(\frac{k+l}{2^p}-\frac{k'}{2^{p'}}\right)^{\frac{1}{2}}}>\psi\left(\frac{2^p}{l}\right)\right)
$$

The inequality follows from the definition of ordering and the fact that the correlation coefficient of two random variables appearing in the last term of (33) is larger than that of the second term. Since $p' \leqq p+2$, we obtain by Lemma 1 and Lemma 5 that

$$
\begin{aligned}
P(E_jE_{j_t}) &\leqq ce^{-\delta\left(\frac{k'-k2^{p'-p}}{2^{p'-p}}\right)\psi^2\left(\frac{2^p}{l}\right)} P(E_j) \\
&\leqq ce^{-\delta'(k'-k2^{p'-p})} P(E_j),
\end{aligned}
\tag{34}
$$

where c, δ and δ' are positive constants. Here we remark that the number of E_{j_t} appearing in the present case is less than $(k'-k2^{p'-p})$ for fixed pair (p', k') because $\frac{l}{2^p} \geqq \frac{l'}{2^{p'}}$. Similarly, for the case of $\frac{k}{2^p} > \frac{k'}{2^{p'}}$ we have

$$
P(E_jE_{j_t}) \leqq ce^{-\delta'(l2^{p'-p}-l')} P(E_j). \tag{35}
$$

Considering the same situation for $\frac{k}{2^p} > \frac{k'}{2^{p'}}$, we have

$$
\begin{aligned}
\Sigma' P(E_jE_{j_t}) &\leqq 2cP(E_j)\sum_{p'=p}^{p+2}\left\{\sum_{k'=k2^{p'-p}}^{(k+l)2^{p'-p}}{}' (k'-k2^{p'-p})e^{-\delta'(k'-k2^{p'-p})}\right. \\
&\qquad\left.+\sum_{l'=1}^{l2^{p'-p}}(l2^{p'-p}-l')e^{-\delta(l2^{p'-p}-l')}\right\} \\
&\leqq \alpha P(E_j),
\end{aligned}
\tag{36}
$$

where α is an absolute constant.

For the computation of $P(E_jE_{j_t})$ where E_{j_t} appears in the summation of Σ'', we apply Lemma 4. Using the same expression for E_t and E_{j_t} as before, for the case of $\frac{k}{2^p} \leqq \frac{k'}{2^{p'}} < \frac{k+l}{2^p} \leqq \frac{k'+l'}{2^{p'}}$, we have

$$
\begin{aligned}
P(E_jE_{j_t}) &\leqq P\left(\frac{X\left(\frac{k+l}{2^p}\right)-X\left(\frac{k}{2^p}\right)}{\left(\frac{l}{2^p}\right)^{\frac{1}{2}}} > \psi\left(\frac{2^p}{l}\right), \frac{X\left(\frac{k+l}{2^p}\right)-X\left(\frac{k'}{2^{p'}}\right)}{\left(\frac{k+l}{2^p}-\frac{k'}{2^{p'}}\right)^{\frac{1}{2}}} > \psi\left(\frac{2^{p'}}{l'}\right)\right) \\
&\leqq ce^{-\delta\psi^2\left(\frac{2^{p'}}{l'}\right)} P(E_j) \\
&\leqq ce^{-\delta'p'} P(E_j)
\end{aligned}
\tag{37}
$$

where c, c', δ and δ' are positive constants. Similarly, for the case of $\frac{k}{2^p} \leqq \frac{k'}{2^{p'}} < \frac{k'+l'}{2^{p'}} \leqq \frac{k+l}{2^p}$, we obtain the same result. Combining all the cases, we have

$$\begin{aligned}\Sigma'' P(E_j E_{j'}) &\leqq cP(E_j) \sum_{p'=p}^{p+5\log p} p'^2 2^{p'-p} l e^{-\delta' p'} \\ &\leqq cP(E_j) \sum_{p'=p}^{p+5\log p} p'^3 e^{-\delta' p'} \qquad (38) \\ &= \beta P(E_j),\end{aligned}$$

where β is an absolute constant. (32), (36) and (38) establish the validity of (iii a). Therefore we may aplly Lemma 2 in our case and Theorem 1 is proved completely. Q. E. D.

Syracuse University, Technion, Israel,
and Shizuoka University.

References

[1] P. Lévy, Théorie de l'addition des variables aléatoires, 1937.
[2] T. Sirao, On the uniform continuity of Wiener process, J. Math. Soc. Japan, 6, (1954), 332-335.
[3] P. Lévy, Le mouvement brownien, Mém. Sci. Math., fasc. 126, 1954.
[4] K. L. Chung and P. Erdös, On the application of the Borel-Cantelli Lemma, Trans. Amer. Math. Soc., 72, (1952), 179-189.

Reprinted from ILLINOIS JOURNAL OF MATHEMATICS
Vol. 4, No. 4, December 1960
Printed in U.S.A.

ON LAST EXIT TIMES

BY
K. L. CHUNG[1]

1. The terminology and notation of this paper follow that of my book [1], where not explicitly explained. Results cited without amplification can also be found in the book.

Let $\{x_t, t \geqq 0\}$ be a well-separable, measurable Markov chain with the discrete state space I, the initial distribution $\{p_i\}$ and stationary, standard transition matrix $((p_{ij}))$, $i, j \in I$. Let

$$(1)\quad {}_ip_{kj}(t) = P\{x(t_0 + t, w) = j, x(t_0 + s, w) \neq i, 0 < s < t \mid x(t_0, w) = k\};$$

for every $t_0 \geqq 0$ for which the conditional probability is defined; thus ${}_ip_{kj}(t) \equiv 0$ if $k = i$ or $i = j$, by stochastic continuity. We note that if k is a stable state and $k = i$, the definition (1) differs from the one adopted in [1].

Writing as usual

$$S_i(w) = \{t : x(t, w) = i\}, \qquad \overline{S_i(w)} = \text{closure of } S_i(w),$$

we define

$$(2)\qquad \gamma_i(t, w) = \sup\{\overline{S_i(w)} \cap [0, t]\}$$

and call it *the last exit time from i before time t*. The separability and measurability of the process ensure that the corresponding w-function $\gamma_i(t)$ is a random variable. Under the hypothesis that $x(0, w) = i$, the stochastic continuity of the process implies that $\gamma_i(t)$ has a distribution function $\Gamma_i(\cdot, t)$ vanishing at zero, continuous in $(0, t)$, and making a jump of magnitude $p_{ii}(t)$ at t to reach the value one. We have clearly, if $0 \leqq s \leqq t$,

$$(3)\qquad \Gamma_i(s, t) = \textstyle\sum_{k \neq i} p_{ik}(s)[1 - F_{ki}(t - s)],$$

where F_{ki} is the *first entrance time distribution from k to i*. We define similarly

$$(4)\qquad \begin{aligned} \Gamma_{ij}(s, t) &= P\{\gamma_i(t, w) \leqq s; x(t, w) = j \mid x(0, w) = i\} \\ &= \textstyle\sum_k p_{ik}(s)\, {}_ip_{kj}(t - s), \end{aligned}$$

noting that the term corresponding to $k = i$ vanishes. Thus we have

$$\Gamma_i(s, t) = \textstyle\sum_{j \neq i} \Gamma_{ij}(s, t).$$

2. The set of sample functions with $x(0, w) = i$ and $x(t, w) = j$ can be decomposed into subsets according to the location of $\gamma_i(t, w)$ in a dyadic partition of $[0, t]$. Since the terminating dyadics $\{v2^{-n}\}$ form a separability set,

Received November 16, 1959.

[1] This research was supported by the United States Air Force through the Air Force Office of Scientific Research and Development Command.

we have by a familiar argument

$$(5)\qquad p_{ij}(t) = \lim_{n\to\infty} \sum_{v=1}^{[2^n t]} p_{ii}\left(\frac{v-1}{2^n}\right) \sum_k p_{ik}\left(\frac{1}{2^n}\right) {}_ip_{kj}\left(t - \frac{v}{2^n}\right).$$

Let us write this more suggestively as

$$(6)\qquad p_{ij}(t) = \lim_{n\to\infty} \int_0^t \phi_{ij}^{(n)}(t-s)\, d\pi_i^{(n)}(s),$$

where

$$\pi_i^{(n)}(s) = \sum_{v=1}^{[2^n s]} p_{ii}\left(\frac{v-1}{2^n}\right)\frac{1}{2^n}, \qquad \phi_{ij}^{(n)}(s) = 2^n \sum_k p_{ik}\left(\frac{1}{2^n}\right) {}_ip_{kj}(s).$$

This motivates the investigation of $\phi_{ij}^{(n)}(s)$ as $n \to \infty$, which we now proceed with. For $i \neq j$, $\delta > 0$, and $s \geqq 0$, we set

$$(7)\qquad \phi_{ij}(\delta; s) = (1/\delta)\textstyle\sum_k p_{ik}(\delta)\, {}_ip_{kj}(s).$$

1°. *Given $\varepsilon > 0$, there exists $\eta(\varepsilon) > 0$ such that for any positive integer m satisfying $m\delta \leqq \eta(\varepsilon)$ we have*

$$(8)\qquad \phi_{ij}(m\delta; s) \geqq (1-\varepsilon)\phi_{ij}(\delta; s)$$

for all $s \geqq 0$.

Proof. Consider the *discrete skeleton* $\{x_{n\delta}, n \geqq 0\}$ and write

$${}_ip_{ik}^{(m-v)} = P\{x((m-v)\delta, w) = k,\ x(n\delta, w) \neq i,\quad 1 \leqq n \leqq m-v-1 \mid x(0,w) = i\}.$$

We have

$$p_{ik}(m\delta) = \sum_{v=0}^{m-1} p_{ii}(v\delta)\, {}_ip_{ik}^{(m-v)}(\delta).$$

This is the *last entrance formula* for the discrete skeleton. Furthermore it follows from the definitions that

$${}_ip_{ik}^{(m-v)}(\delta) = \textstyle\sum_{l\neq i} p_{il}(\delta)\, {}_ip_{lk}^{(m-v-1)}(\delta) \geqq \sum_l p_{il}(\delta)\, {}_ip_{lk}((m-v-1)\delta).$$

Hence we have, using the semigroup property of $(({}_ip_{lj}))$ for a fixed i,

$$\begin{aligned}\phi_{ij}(m\delta; s) &\geqq \frac{1}{m\delta}\sum_{v=0}^{m-1} p_{ii}(v\delta) \sum_k \sum_l p_{il}(\delta)\, {}_ip_{lk}((m-v-1)\delta)\, {}_ip_{kj}(s)\\ &= \frac{1}{m\delta}\sum_{v=0}^{m-1} p_{ii}(v\delta) \sum_l p_{il}(\delta)\, {}_ip_{lj}((m-v-1)\delta + s)\\ &\geqq \frac{1}{m\delta}\sum_{v=0}^{m-1} p_{ii}(v\delta) \sum_l p_{il}(\delta)\, {}_ip_{lj}(s)\, {}_ip_{jj}((m-v-1)\delta).\end{aligned}$$

We choose $\eta(\varepsilon)$ so that if $m\delta \leqq \eta(\varepsilon)$, then

$$\min_{0\leqq v<m} p_{ii}(v\delta) \cdot \min_{0\leqq v<m} {}_ip_{jj}((m-v-1)\delta) \geqq 1-\varepsilon.$$

This is possible since $\lim_{t\downarrow 0}\, {}_jp_{jj}(t) = 1$. Then we have

$$\phi_{ij}(m\delta; s) \geqq \frac{1}{m\delta}\, m(1-\varepsilon) \sum_l p_{il}(\delta)\, {}_jp_{lj}(s) = (1-\varepsilon)\phi_{ij}(\delta; s).$$

2°. *For every $i \neq j$ and every $s \geqq 0$,*

$$\lim_{\delta\downarrow 0}\phi_{ij}(\delta; s)$$

exists and is a bounded function of s in any finite interval.

Proof. Let

$$\liminf_{\delta\downarrow 0}\phi_{ij}(\delta; s) = g_{ij}(s).$$

Since the series in (7) is dominated by the series $\sum_k p_{ik}(\delta)$ which converges uniformly with respect to δ in any finite interval, we see that $\phi_{ij}(\delta; s)$ is continuous in δ for each s. Furthermore, it follows from (8) that g_{ij} is bounded in any finite interval since $\phi_{ij}(\eta, \cdot\,)$ is for every η. Now for all sufficiently small δ we may choose $m\delta$ so that $\phi_{ij}(m\delta; s)$ is near $g_{ij}(s)$; hence the existence of the limit asserted in 2° follows from the inequality (8) and the definition of $g_{ij}(s)$. The above argument is similar to one by Kolmogorov [4] (cf. Theorem 2.5 of [1]); indeed Kolmogorov's theorem corresponds to the case $s = 0$ here.

If we could show that the convergence in 2° is uniform with respect to s, or equivalently (see below in 3°) that the limit function g_{ij} is continuous, then we could pass to the limit in (6) and obtain the desired result. Unfortunately we have to do this in a rather devious way. Let us denote by $P_{ij}(t)$, ${}_jP_{kj}(t)$, $G_{ij}(t)$, and $\Phi_{ij}^{(n)}(t)$ the integrals of p_{ij}, ${}_jp_{kj}$, g_{ij}, and $\phi_{ij}^{(n)}$ from 0 to t.

3°. *We have*

$$P_{ij}(t) = \int_0^t G_{ij}(t-s)p_{ii}(s)\, ds. \tag{9}$$

Proof. We note first the following complement to (5) or (6):

$$p_{ij}(t) \geqq \int_0^t \phi_{ij}^{(n)}(t-s)\, d\pi_i^{(n)}(s) \tag{10}$$

which is immediate by the sample function interpretation. We can therefore integrate (6) under the limit sign by dominated convergence and obtain

$$P_{ij}(t) = \lim_{n\to\infty}\int_0^t \Phi_{ij}^{(n)}(t-s)\, d\pi_i^{(n)}(s). \tag{11}$$

It follows from (8) that given any $\varepsilon > 0$, there exists an $n_0(\varepsilon)$ independent of s such that for all $n_2 > n_1 > n_0(\varepsilon)$ we have

$$\Phi_{ij}^{(n_1)}(s) \geqq (1-\varepsilon)\Phi_{ij}^{(n_2)}(s) \tag{12}$$

for all $s \geqq 0$. The $\Phi_{ij}^{(n)}$ and G_{ij} are continuous functions and $\Phi_{ij}^{(n)}$ converges to G_{ij} by 1° and 2°. These are the hypotheses in Dini's theorem on uniform

convergence except that the condition of monotonicity is weakened to (12). The usual proof of Dini's theorem carries over without ado, establishing that the convergence of $\Phi_{ij}^{(n)}(s)$ is uniform with respect to s in $[0, t]$ for any finite t. Now simultaneous passage to the limit of the integrand and integrator in (11) is permitted, and (9) follows since

$$\lim_{n\to\infty} \pi_i^{(n)}(t) = \int_0^t p_{ii}(s)\, ds.$$

4°. *The G_{ij}'s satisfy the following system of equations:*

$$G_{ij}(s+t) - G_{ij}(t) = \sum_k G_{ik}(s)\, {}_ip_{kj}(t) \tag{13}$$

for all $s \geqq 0$, $t \geqq 0$.

Proof. Integrating (10) we see that

$$P_{ij}(t) \geqq \int_0^t \Phi_{ij}^{(n)}(t-s)\, d\pi_i^{(n)}(s) \geqq \Phi_{ij}^{(n)}(s)\pi_i^{(n)}(t-s)$$

for any $s \in (0, t)$. Since $\sum_j P_{ij}(t) = t$, it follows that $\sum_j \Phi_{ij}^{(n)}(s)$ converges uniformly with respect to n (and also with respect to s in any finite interval). We have by definition

$$\sum_k \Phi_{ik}^{(n)}(s)\, {}_ip_{kj}(t) = \int_0^s \sum_k \phi_{ik}^{(n)}(u)\, {}_ip_{kj}(t)\, du$$

$$= \int_0^s \phi_{ij}^{(n)}(u+t)\, du = \Phi_{ij}^{(n)}(s+t) - \Phi_{ij}^{(n)}(t).$$

Letting $n \to \infty$, we obtain (13) on account of the stated uniform convergence.

5°. *The G_{ij}'s have continuous derivatives in $[0, \infty)$ satisfying*

$$G'_{ij}(s+t) = \sum_k G'_{ik}(s)\, {}_ip_{kj}(t) \tag{14}$$

for all $s > 0$, $t > 0$.

Proof. An application of Fubini's theorem on differentiation to the system (13) yields the system (14) for each $t > 0$ and almost all s (depending on t). Hence it also holds if $s \notin Z$ and $t \notin Z_s$ where Z and Z_s are sets of measure zero. If we consider G'_{ij} ($\geqq 0$) as the right-hand lower derivate, then we see directly from (13), upon taking the proper difference quotients and using Fatou's lemma, that the inequality obtained from (14) by changing " = " into " $\geqq$ " holds for all positive s and t. Fix an $s \notin Z$; if the left member of (14) is strictly greater than the right member for a certain value of t, then the same is true for all greater values of t by the semigroup property of $(({}_ip_{kj}))$. This being impossible by a previous assertion, the equation (14) is true for $s \notin Z$ and all $t > 0$. Now for arbitrary positive s and t write $s + t = s' + t'$ where $0 < s' < s$, $s' \notin Z$. Applying what we have just proved to $G'_{ij}(s' + t')$ and using the semigroup property again, we see that (14) is true for all positive s and t. The continuity of G'_{ij} is then a consequence of the system (14); see

Theorem 2.3 of [1]. Finally, this implies the continuous differentiability of G_{ij} by another well-known theorem of Dini.

We are now ready to state the following result.

THEOREM 1. *If $i \neq j$, then*

$$p_{ij}(t) = \int_0^t p_{ii}(s) g_{ij}(t - s)\, ds, \qquad 0 \leqq t < \infty, \tag{15}$$

where g_{ij} is the limit of $\phi_{ij}(\delta; s)$ as $\delta \downarrow 0$, uniformly with respect to s in any finite interval. The system (14) *holds with $G'_{ij} = g_{ij}$; and*

$$g_i(t) = \textstyle\sum_{j \neq i} g_{ij}(t)$$

is continuous and nonincreasing for $t > 0$. We have $g_{ij}(0+) = p'_{ij}(0)$ and $g_i(0+) = -p'_{ii}(0)$.

Proof. It is permissible to differentiate (9) under the integral sign; the result is (15) with g_{ij} replaced by G'_{ij}. Using this result, we have by (14)

$$\begin{aligned}\phi_{ij}(\delta; s) &= \frac{1}{\delta} \sum_k \int_0^\delta p_{ii}(\delta - u) G'_{ik}(u)\; {}_ip_{kj}(s)\, ds \\ &= \frac{1}{\delta} \int_0^\delta p_{ii}(\delta - u) G'_{ij}(u + s)\, ds.\end{aligned}$$

When $\delta \downarrow 0$, the left member tends to $g_{ij}(s)$, and the right member tends uniformly to $G'_{ij}(s)$ by continuity. Hence $g_{ij} \equiv G'_{ij}$, and (15) is proved. It now follows from (14) that if $g_i(s) < \infty$, then for all $t > 0$,

$$g_i(s + t) = \textstyle\sum_{k \neq i} g_{ik}(s)[1 - F_{ki}(t)] \leqq g_i(s) < \infty.$$

Summing (15) over $j \neq i$ we see that $g_i(s) < \infty$ for a.a. s, hence indeed for all $s > 0$; and furthermore g_i is continuous there by the equation above, since all F_{ki} are. It is not difficult to show that g_i is absolutely continuous and that for each s and a.a. t we have

$$g'_i(s + t) = -\textstyle\sum_{k \neq i} g_{ik}(s) f_{ki}(t),$$

where $f_{ki} = F'_{ki}$. The last assertion of the theorem follows from (15).

Theorem 1 has been proved by Jurkat [2]. His treatment is algebraic-analytical and does not require the "row condition" that $\sum_j p_{ij}(t) = 1$ for the transition matrix. The above proof is new and shows more directly the relation to the "movement" of the Markov chain. The probabilistic significance becomes clearer in the next statement.

THEOREM 2. *For each t, $\Gamma_{ij}(s, t)$ as defined in* (4) *has a continuous derivative with respect to s given by $p_{ii}(s) g_{ij}(t - s)$; and the distribution $\Gamma_i(s, t)$ as defined in* (3) *has a continuous density in s given by $p_{ii}(s) g_i(t - s)$. One version of the conditional probability*

$$P\{x(t, w) = j \mid x(0, w) = i, \gamma_i(t, w) = s\}$$

is equal to $g_{ij}(t-s)/g_i(t-s)$; *thus as a function of* (s, t) *it is a function of the difference* $t - s$ *only.*

Proof. We have by substituting from (15) and using (14),

$$\Gamma_{ij}(s, t) = \sum_{k \neq i} \int_0^s p_{ii}(u) g_{ik}(s-u)\, {}_i p_{kj}(t-s)\, du$$
$$= \int_0^s p_{ii}(u) g_{ij}(t-u)\, du.$$

Summing over $j \neq i$ we have

$$\Gamma_i(s, t) = \int_0^s p_{ii}(u) g_i(t-u)\, du.$$

These formulas establish the first two assertions of the theorem. The last assertion follows from (15) written in the form

$$p_{ij}(t) = \int_0^t \frac{g_{ij}(t-s)}{g_i(t-s)}\, d_s \Gamma_i(s, t).$$

The dual of Theorem 1 is well known. We give it here for the sake of comparison.

THEOREM 3. *If* $i \neq j$, *then*

$$p_{ij}(t) = \int_0^t p_{jj}(t-s) f_{ij}(s)\, ds, \qquad 0 \leqq t < \infty, \tag{16}$$

where f_{ij} *is the continuous derivative (density) of the first entrance time distribution* F_{ij}. *We have*

$$f_{ij}(s+t) = \sum_k {}_j p_{ik}(s) f_{kj}(t), \qquad s > 0,\ t > 0. \tag{17}$$

Proof. That

$$p_{ij}(t) = \int_0^t p_{jj}(t-s)\, dF_{ij}(s) \tag{18}$$

is a special case of Theorem II. 11.8 of [1]; that F_{ij} has a continuous derivative f_{ij} satisfying (17) can be shown (oral communication by D. G. Austin) by differentiating the following identity:

$$1 - F_{ij}(s+t) = \sum_k {}_j p_{ik}(s)[1 - F_{kj}(t)] \tag{19}$$

as in 5° above.

3. Since the basic formula (18) can be proved by a probabilistic argument relying on a special case (where the optional time is the first entrance time) of the strong Markov property, it is natural to ask if this genre of reasoning can also be dualized to yield a proof of Theorem 1, at least in the form corresponding to (18). This will now be shown.

For a given $T > 0$, we define the *reversed Markov chain in* $[0, T]$ as follows:

$$x_T^*(t, w) = x(T - t, w), \qquad 0 \leqq t \leqq T. \tag{20}$$

This has the state space I and the *nonstationary* transition probabilities given by

$$p_T^*(s,t;j,i) = P\{x_T^*(t,w) = i \mid x_T^*(s,w) = j\} = \frac{p_i(T-t)}{p_j(T-s)} p_{ij}(t-s), \tag{21}$$

where

$$p_i(t) = \textstyle\sum_k p_k\, p_{ki}(t)$$

is the *absolute distribution of* x_t. The *first entrance time distribution from* j *to* i *and starting at time* s is given by

$$\begin{aligned} F_T^*(s, t; j, i) &= 1 - P\{x_T^*(u, w) \neq i, s \leqq u \leqq t \mid x_T^*(s, w) = j\} \\ &= 1 - P\{x(T-u, w) \neq i, s \leqq u \leqq t \mid x(T-s, w) = j\} \\ &= 1 - \frac{1}{p_j(T-s)} \sum_k p_k(T-t)\; {}_ip_{kj}(t-s) \\ &= 1 - \frac{1}{p_j(T-s)} \sum_k p_k \Gamma_{kj}(T-t, T-s). \end{aligned} \tag{22}$$

We shall prove the following first entrance formula for a Markov chain with nonstationary transition probabilities, corresponding to (18).

THEOREM 4.[2] *Let* $\{y_t, 0 \leqq t \leqq T\}$ *be a measurable Markov chain with the state space* I *and transition probability function*

$$p(s, t; i, j) = P\{y(t, w) = j \mid y(s, w) = i\}, \qquad 0 \leqq s \leqq t \leqq T, i, j \in I.$$

The following assumptions are made:

(i) *for fixed* t, i, *and* j, $p(s, t; i, j)$ *is right continuous in* $s \leqq t$;

(ii) *there exist a set* Ω *of probability one and a denumerable dense set* $R \in [0, T]$ *such that if* $w \in \Omega$, *then for all* $t \in [0, T)$, $y(t, w)$ *is a limiting value of* $y(r, w)$ *as* $r \downarrow t$, $r \in R$.

(iii) *for each* t *and* j, $P\{t \in \overline{S_j(w)} - S_j(w)\} = 0$.

We have then if $i \neq j$, $0 \leqq s < t < T$,

$$p(s, t; i, j) = \int_s^t p(u, t; j, j)\, d_u F(s, u; i, j), \tag{23}$$

where

$$F(s, u; i, j) = P\{y(t, w) = j \quad \textit{for some} \quad t \in [s, u] \mid y(s, w) = i\}.$$

Remark. The condition (ii) may be roughly described as *"right separable with respect to* R." Any process has such a version.

[2] This theorem can be easily modified to yield the strong Markov property for the process. While this property in the nonstationary case has been discussed by other authors, I was unable to find a result which would cover the situation here.

Proof. Let R be enumerated as $\{r_m\}$, and let $\{r_m, 1 \leqq m \leqq n\}$ be ordered as $r_1^{(n)} < \cdots < r_n^{(n)}$. Define

$$\beta(s, w) = \inf\{t : t > s, y(t, w) = j\},$$
$$\beta^{(n)}(s, w) = \inf\{r_m^{(n)} : r_m^{(n)} > \beta(s, w); y(r_m^{(n)}, w) = j\}.$$

Clearly, β and $\beta^{(n)}$ are optional random variables (see [1] for the definition which is also valid in the nonstationary case). We shall prove that for almost all w in the set $\{w : \beta(s, w) \leqq t\}$, we have

$$P\{y(t, w) = k \mid \beta(s, w)\} = p(\beta(s, w), t; j, k) \tag{24}$$

for all k. This implies (23) when $k = j$.

To prove (24) we have only to repeat the argument in the stationary case used in [1]. It follows from (ii) that $\beta^{(n)}(s, w) \downarrow \beta(s, w)$ for almost all w. Hence if $s \leqq t' \leqq t$,

$$\begin{aligned}
P\{y(t, w) = k; \beta(s, w) < t'\} &= \lim_{n\to\infty} P\{y(t, w) = k; \beta^{(n)}(s, w) < t'\} \\
&= \lim_{n\to\infty} \sum_{r_m^{(n)}<t'} P\{\beta^{(n)}(s, w) = r_m^{(n)}\} P\{y(t, w) = k \mid y(r_m^{(n)}, w) = j\} \\
&= \lim_{n\to\infty} \sum_{r_m^{(n)}<t'} P\{\beta^{(n)}(s, w) = r_m^{(n)}\} p(r_m^{(n)}, t; j, k) \\
&= \lim_{n\to\infty} \int_{\{w:\beta^{(n)}(s,w)<t'\}} p(\beta^{(n)}(s, w), t; j, k) P(dw) \\
&= \int_{\{w:\beta(s,w)<t'\}} p(\beta(s, w), t; j, k) P(dw).
\end{aligned}$$

The truth of this for all $t' \leqq t$ is equivalent to (24), in view of (iii).

We now apply Theorem 4 to the reversed Markov chain $\{x_T^*(t), 0 \leqq t \leqq T\}$ defined in (20). A glance at (21) shows that condition (i) is satisfied. As for condition (ii) we need only take the version of $\{x_t, 0 \leqq t < \infty\}$, denoted by $\{x_-(t), 0 \leqq t < \infty\}$ in [1], which has the property that

$$x_-(t, w) = \liminf_{r\uparrow t, r\in R} x(r, w), \qquad 0 < t < \infty;$$

then we have

$$x_T^*(t, w) = \liminf_{r\downarrow t, r\in R_T} x_T^*(r, w)$$

where R_T consists of the numbers $T - r$ where $r \in R$, $0 \leqq r \leqq T$. Thus (ii) is satisfied. It is known that (iii) is true for $\{x_-(t)\}$, hence also for $\{x_T^*(t)\}$.

We take $p_i = 1$ for $\{x_t, t \geqq 0\}$ so that $p_k(t) \equiv p_{ik}(t)$ for all k. Applying (23) with $s = 0$, interchanging i and j, and substituting from (21) and (22), we obtain after a trivial simplification:

$$p_{ij}(t) = \int_0^t p_{ii}(t - u) \frac{-d_u \Gamma_{ij}(T - u, T)}{p_{ii}(T - u)} = \int_0^t p_{ii}(t - u)\, dG^T(u), \tag{25}$$

say. Without using the results in §2, we can proceed as follows. Each G^T is a nondecreasing continuous function in $[0, T]$ which may be normalized by making $G^T(0) = 0$. For each positive integer n, and all $T \geqq n$, we have from (25)

$$G^T(n) \leqq \frac{p_{ij}(n)}{\min_{0\leqq t\leqq n} p_{ii}(t)} < \infty.$$

Applying Helly's selection principle first in each $[0, n]$, and then diagonalizing, we see that there exist a subsequence $G^{(n_m)}$, and a G_{ij} nondecreasing in $[0, \infty)$ and bounded in every finite interval, such that

$$\lim_{m\to\infty} G^{(n_m)}(u) = G(u), \qquad 0 \leqq u < \infty.$$

Passing to the limit in (25), we obtain

$$p_{ij}(t) = \int_0^t p_{ii}(t-u)\, dG_{ij}(u), \qquad 0 \leqq t < \infty, \tag{26}$$

for which it now appears that G_{ij} must be continuous since p_{ij} and p_{ii} are. It is possible to start from (26) and derive further properties of G_{ij}.

However, we shall look back at Theorem 2 and observe at once that

$$\frac{-1}{p_{ii}(T-u)} \frac{\partial}{\partial u} \Gamma_{ij}(T-u, T) = g_{ij}(u).$$

Hence G^T is actually independent of T, and we have identified (26) with (15). For given p_{ij} and p_{ii}, the uniqueness of the G_{ij} in (26) is of course a known fact, as is best seen by taking Laplace transforms. Finally, we remark that the generalization of (19) to the reversed Markov chain leads to the equation (13).

4. In an interesting special case the treatment of §3 can be simplified. This case in rather restrictive for the purposes of this paper, but there a "*duality principle*" holds which simplifies the preceding considerations.

Consider a *class* C of mutually communicating states containing two distinct states 1 and 2. We set

$$e_i = \int_0^\infty {}_2p_{1i}(t)\, dt + \int_0^\infty {}_1p_{2i}(t)\, dt, \qquad i \in C. \tag{27}$$

It can be shown that $0 < e_i < \infty$ for every i, and

$$\sum_j e_j\, p_{ji}(t) \leqq e_i, \qquad i \in C. \tag{28}$$

This is proved in [1; Theorem II. 13.5] for a recurrent class with the inequality strengthened into an equality; for an arbitrary class the proof requires only an obvious change which necessitates the inequality in general. The existence of a positive solution $\{e_i\}$ for the system of inequalities (28) has been shown independently by D. G. Kendall [3] without the explicit formula (27). It is

known that equality holds in (28) for all i if and only if the class is recurrent, and then there is a unique positive solution apart from a constant factor. In a recurrent-positive class we may take $e_i = \lim_{t\to\infty} p_{ii}(t)$; then $\sum_{i\epsilon C} e_i = 1$, and $\{e_i\}$ yields the stationary distribution of the Markov chain (on C). If this is taken to be the initial distribution, then the reversed chain in §3 will have stationary transition probabilities. The case under discussion here is more general in the sense that $\{e_i\}$ plays the role of a stationary pseudo-distribution even though $\sum_{i\epsilon C} e_i$ may diverge.

We define the *dual matrix* $((p_{ij}^*))$ as follows:

$$p_{ij}^*(t) = (e_j/e_i)p_{ji}(t), \qquad i, j \epsilon C; \quad t \geqq 0. \tag{29}$$

The Markov (semigroup) property of the dual matrix is at once verified, but (28) shows that it is only substochastic. We may however make it stochastic by the usual device of adjoining a new state θ and setting

$$p_{i\theta}^*(t) = 1 - \sum_{j\epsilon C} p_{ij}^*(t); \qquad p_{\theta\theta}^*(t) \equiv 1; \qquad p_{\theta j}^*(t) \equiv 0, \quad j \epsilon C.$$

Thus enlarged, the dual matrix becomes a stochastic transition matrix with which we may associate a Markov chain[3] $\{x_t^*, t \geqq 0\}$ with state space $C^* = C \cup \{\theta\}$ such that

$$p_{ij}^*(t) = P\{x^*(t, w) = j \mid x^*(0, w) = i\}.$$

Taking a well-separable and measurable version of the *dual chain* $\{x_t^*, t \geqq 0\}$, we can introduce the *taboo probabilities*

$$_Hp_{ij}^*(t) = P\{x^*(t, w) = j; x^*(s, w) \not\epsilon H, 0 < s < t \mid x^*(0, w) = i\}, \tag{30}$$

where H is an arbitrary subset of C. Similarly,

$$F_{ij}^*(t) = P\{x^*(s, w) = j \quad \text{for some} \quad s \epsilon [0, t] \mid x^*(0, w) = i\}.$$

Now we have, generalizing (29),

$$_Hp_{ij}^*(t) = (e_j/e_i)\ _Hp_{ji}(t). \tag{31}$$

This follows from an analytic way of defining the probability in (30); we have in fact

$$_Hp_{ij}^*(t) = \lim_{n\to\infty} \sum_{i_1 \not\epsilon H} \sum_{i_2 \not\epsilon H} \cdots \sum_{i_{n-1} \not\epsilon H} p_{ii_1}^*(t/n)p_{i_1i_2}^*(t/n) \cdots p_{i_{n-1}j}^*(t/n).$$

Hence (31) follows from (29).

Let us write the first entrance formula for the dual chain corresponding to (16):

$$p_{ji}^*(t) = \int_0^t p_{ii}^*(t - s)f_{ji}^*(s)\, ds, \qquad i, j \epsilon C,$$

where f_{ji}^* is the continuous derivative of F_{ji}^*. Substituting from (29) we ob-

[3] This is not to be confused with the $x_T^*(t)$ in §3.

tain (15) with

$$g_{ij}(s) = (e_j/e_i)\, f^*_{ji}(s). \tag{32}$$

The formula corresponding to (17) is

$$f^*_{ji}(s + t) = \sum_k {}_i p^*_{jk}(s) f^*_{ki}(t).$$

By (31) and (32) this is equivalent to

$$(e_i/e_j) g_{ij}(s + t) = \sum_k (e_k/e_j)\, {}_i p_{kj}(s)(e_i/e_k) g_{ik}(t)$$

which is (14).

While the above treatment is not adequate for the general results in §§2–3, it seems worthwhile mentioning that it is already sufficient as a basis for a derivation of Ornstein's differentiability theorem [5]. Given the state i, consider the class C in which i belongs. Applying (14) to all $j \in C - \{i\}$, for which the method of this section suffices, and summing, we obtain

$$p_{ii}(t) + \int_0^t p_{ii}(t - s) c_i(s)\, ds = d_i(t) \tag{33}$$

where $c_i = \sum_{j \in C - \{i\}} g_{ij}$ is a continuous function and $d_i = \sum_{j \in C} p_{ij}$ is a non-decreasing, continuously differentiable function. The integral equation (33) for p_{ii} may be used as the starting point to prove the continuous differentiability of p_{ii}; see [2] or [1].

References

1. K. L. Chung, *Markov chains with stationary transition probabilities*, Heidelberg, 1960.
2. W. B. Jurkat, *On the analytic structure of semi-groups of positive matrices*, Math. Zeitschrift, vol. 73 (1960), pp. 346–365.
3. D. G. Kendall, *Unitary dilations of one-parameter semigroups of Markov transition operators, and the corresponding integral representations for Markov processes with a countable infinity of states*, Proc. London Math. Soc. (3), vol. 9 (1959), pp. 417–431.
4. A. N. Kolmogorov, *On some problems concerning the differentiability of the transition probabilities in a temporally homogeneous Markov process having a denumerable number of states*, Moskov. Gos. Univ. Učenye Zapiski. Mat., vol. 148 (4) (1951), pp. 53-59 (in Russian).
5. D. S. Ornstein, *The differentiability of transition functions*, Bull. Amer. Math. Soc., vol. 66 (1960), pp. 36–39.

Syracuse University
Syracuse, New York

Reprinted from Illinois Journal of Mathematics
Vol. 5, No. 3, September 1961
Printed in U.S.A.

SOME REMARKS ON TABOO PROBABILITIES

BY

K. L. Chung[1]

Consider a discrete-parameter homogeneous Markov chain $\{x_n, n \geq 0\}$ with state space I and one-step transition matrix $((p_{ij}))$, $i, j \in$ I. For any subset H of I we define the taboo probability

$$_Hp_{ij}^{(n)} = \mathbf{P}\{x_n = j; x_\nu \notin H, 0 < \nu < n \mid x_0 = i\}, \qquad n \geq 1,$$

and set

$$_Hp_{ij}^{*} = \sum_{n=1}^{\infty} {}_Hp_{ij}^{(n)}.$$

When $H = \{k\}$ we write $_kp_{ij}^{(n)}$ for $_Hp_{ij}^{(n)}$. Furthermore we write $f_{ij}^{(n)}$ for $_jp_{ij}^{(n)}$ and $e_{ij}^{(n)}$ for $_ip_{ij}^{(n)}$. Thus,

$$f_{ij}^{*} = \sum_{n=1}^{\infty} f_{ij}^{(n)}, \qquad e_{ij}^{*} = \sum_{n=1}^{\infty} e_{ij}^{(n)}.$$

The quantity f_{ij}^* is familiar, the quantity e_{ij}^* has been studied in [1] under the notation e_{ij}. We set also

$$m_{ij} = \sum_{n=1}^{\infty} n f_{ij}^{(n)};$$

and $_Hp_{ij}^{(0)} = \delta_{ij}$.

In this paper a "well known" statement means one which can be found in [1], particularly §I.9 there which treats taboo probabilities. What follows may be regarded as some interesting corollaries of well-known results which seem worth stating. They are engendered by a generalization (Proposition 7) of a recent result of Spitzer [3]. This will be placed where it properly belongs, and the proof will be strictly elementary. In doing so we shall define a new binary relation between the states of a Markov chain.

Recall that two states i and j belong to the same recurrent class if and only if $f_{ij}^* = f_{ji}^* = 1$, or alternatively $f_{ii}^* = 1$ and $f_{ij}^* > 0$. A subset C of I is said to form an *equitable class* if and only if for every i and j in C we have

$$e_{ij}^* = 1. \tag{1}$$

A well-known example of an equitable class is the following: x_n is the sum of n independent and identically distributed random variables with mean zero, or more generally, x_n is a recurrent Markov chain with stationary and independent increments.

We have the following characterization.

Proposition 1. *An equitable class C is recurrent. A recurrent class is equitable if and only if for each i in C, we have*

$$\sum_{j \in C} p_{ji} = 1, \tag{2}$$

namely when $((p_{ij}))$ restricted to C is doubly stochastic.

Received October 6, 1960.

[1] This work is supported in part by the Office of Scientific Research of the United States Air Force.

Proof. The first assertion is trivial since $f_{ii}^* = e_{ii}^* = 1$, and $e_{ij}^* = 1$ implies $f_{ij}^* > 0$. Hence the second criterion for recurrence cited above applies. Next, by a well-known theorem of Derman, the following system of equations

$$u_i = \sum_{j \epsilon C} u_j p_{ji}, \qquad i \epsilon C,$$

has the unique nonnegative solutions $\{u_j, j \epsilon C\}$ given by

$$u_j = c_i e_{ij}^*$$

for an arbitrary $i \epsilon C$ and an arbitrary $c_i \geqq 0$. This and the fact that $e_{ii}^* = 1$ establish the second assertion of the proposition.

PROPOSITION 2. *An equitable class C is positive-recurrent if and only if C is a finite set and*

$$m_{ii} = c \tag{3}$$

for each $i \epsilon C$, where c is the cardinal of C.

Proof. It is well known that the recurrent class C is positive or null according as $\sum_{j \epsilon C} e_{ij}^* < \infty$ or $= \infty$. In the former case it is also well known that

$$e_{ij}^* = m_{jj}/m_{ii} \quad \text{and} \quad \sum_{i \epsilon C} (1/m_{ii}) = 1.$$

The proposition follows from these facts.

PROPOSITION 3. *For a class C to be equitable it is necessary and sufficient that there exists an i in C such that we have $e_{ij}^* = 1$ for each j in C.*

Proof. Necessity is just a part of the definition of equitability. To prove sufficiency we observe that C must be recurrent since i is. Now it is well known that in a recurrent class we have

$$e_{ij}^* e_{jk}^* = e_{ik}^*. \tag{4}$$

It follows that $e_{jk}^* = 1$ for every j and k, proving that C is equitable.

Two distinct states i and j are said to be an *equitable pair* if and only if

$$e_{ij}^* = e_{ji}^* = f_{ij}^* = f_{ji}^* = 1.$$

It is important to note that an equitable pair of states does not necessarily belong to an equitable class, since the latter need not exist.

PROPOSITION 4. *If i and j are an equitable pair, then they belong to the same recurrent class C, and $e_{ik}^* = e_{jk}^*$ for each k in C. Furthermore we have*

$$_jp_{ii}^* = {}_ip_{jj}^*, \qquad {}_if_{ij}^* = {}_jf_{ji}^*,$$

where ${}_if_{ij}^* = {}_Hp_{ij}^*$, ${}_jf_{ji}^* = {}_Hp_{ji}^*$ with $H = \{i, j\}$.

Proof. The first assertion follows from the first criterion of recurrence cited before; the second from (4). The rest follows from the well-known

relations:

$$\frac{f_{ij}^*}{e_{ij}^*} = \frac{1 + {}_jp_{ii}^*}{1 + {}_ip_{jj}^*} = \frac{f_{ij}^*\,{}_jf_{ji}^*}{f_{ji}^*\,{}_if_{ij}^*}\left(= \lim_{n\to\infty} \frac{\sum_{\nu=1}^n p_{ii}^{(\nu)}}{\sum_{\nu=1}^n p_{jj}^{(\nu)}}\right).$$

(The one in the parenthesis is not needed to prove the others, but is the heart of the matter.)

The next proposition, in some form, has been mentioned to me by several persons, including Hoeffding, Orey, and C. J. Stone.

PROPOSITION 5. *Whatever the states i and j, we have*

$$(5) \qquad 0 \leqq \sum_{n=0}^N [p_{ii}^{(n)} - p_{ji}^{(n)}] \leqq \sum_{n=0}^N {}_jp_{ii}^{(n)}, \qquad N \geqq 0.$$

If $e_{ij}^* \leqq 1$ and $e_{ji}^* \leqq 1$, then

$$(6) \qquad 0 \leqq \sum_{n=0}^N [p_{jj}^{(n)} - p_{ji}^{(n)}] \leqq \sum_{n=0}^N {}_ip_{jj}^{(n)}, \qquad N \geqq 0.$$

Remark. The "silent" condition for (5) is of course $f_{ij}^* \leqq 1$ and $f_{ji}^* \leqq 1$. Observe also that if i and j are recurrent, and $e_{ij}^* \leqq 1$, $e_{ji}^* \leqq 1$, then i and j are in fact an equitable pair since $1 = e_{ii}^* = e_{ij}^*e_{ji}^*$.

Proof. We prove (6) only, since (5) is similar and will not be used. We may suppose that $i \neq j$ in (6). By a well-known formula,

$$p_{jj}^{(n)} = {}_ip_{jj}^{(n)} + \sum_{\nu=1}^n p_{ji}^{(\nu)} e_{ij}^{(n-\nu)};$$

consequently

$$\sum_{n=0}^N p_{jj}^{(n)} = \sum_{n=0}^N {}_ip_{jj}^{(n)} + \sum_{\nu=1}^N p_{ji}^{(\nu)} \sum_{n=1}^{N-\nu} e_{ij}^{(n)}$$
$$\leqq \sum_{n=0}^N {}_ip_{jj}^{(n)} + \sum_{\nu=1}^N p_{ji}^{(\nu)},$$

since $\sum_{n=1}^{N-\nu} e_{ij}^{(n)} \leqq e_{ij}^* \leqq 1$. On the other hand, we have by another well-known formula:

$$p_{ji}^{(n)} = \sum_{\nu=0}^n p_{jj}^{(\nu)} e_{ji}^{(n-\nu)},$$

$$\sum_{n=0}^N p_{ji}^{(n)} = \sum_{\nu=0}^N p_{jj}^{(\nu)} \sum_{n=0}^{N-\nu} e_{ji}^{(n)} \leqq \sum_{\nu=0}^N p_{jj}^{(\nu)}$$

since $\sum_{n=0}^{N-\nu} e_{ji}^{(n)} \leqq e_{ji}^* \leqq 1$. These inequalities establish (6).

The next proposition has nothing to do with probability; it is stated here without (trivial) proof for the sake of explicitness.

PROPOSITION 6. *If $\sum_\nu |a_\nu| < \infty$, $|b_n| \leqq A < \infty$ for all n, and $\lim_{n\to\infty}(b_n - b_{n-1}) = 0$, then*

$$\lim_{n\to\infty} \sum_{\nu=1}^n |a_\nu|\,|b_n - b_{n-\nu}| = 0.$$

PROPOSITON 7. *If i and j are an equitable pair and belong to either a null-recurrent class or an aperiodic positive-recurrent class, then we have*

$$(7) \qquad \sum_{n=0}^\infty \{p_{ii}^{(n)} + p_{jj}^{(n)} - p_{ij}^{(n)} - p_{ji}^{(n)}\} = 1 + {}_jp_{ii}^*.$$

Remark 1. It is well known that $1 + {}_jp_{ii}^* = ({}_jf_{ji}^*)^{-1} < \infty$ in a recurrent

class. Since the left member of (7) is symmetric in i and j, (7) implies a partial proof of the second sentence in Proposition 4.

Remark 2. The proposition is not true if i and j belong to a periodic, positive-recurrent class; let $p_{ij} = 1 = p_{ji}$.

Proof. Write the trivial equations:

$$(8) \qquad \sum_{n=1}^{N} p_{ji}^{(n)} = \sum_{\nu=1}^{N} f_{ij}^{(\nu)} \sum_{n=1}^{N} p_{ji}^{(n)} + \sum_{\nu=N+1}^{\infty} f_{ij}^{(\nu)} \sum_{n=1}^{N} p_{ji}^{(n)},$$

$$(9) \qquad \sum_{n=0}^{N} p_{jj}^{(n)} = \sum_{\nu=1}^{N} f_{ij}^{(\nu)} \sum_{n=0}^{N} p_{jj}^{(n)} + \sum_{\nu=N+1}^{\infty} f_{ij}^{(\nu)} \sum_{n=0}^{N} p_{jj}^{(n)}.$$

By well-known formulas, we have

$$(10) \qquad \sum_{n=1}^{N} p_{ij}^{(n)} = \sum_{\nu=1}^{N} f_{ij}^{(\nu)} \sum_{n=0}^{N-\nu} p_{jj}^{(n)},$$

$$(11) \qquad \sum_{n=0}^{N} p_{ii}^{(n)} = \sum_{\nu=1}^{N} f_{ij}^{(\nu)} \sum_{n=1}^{N-\nu} p_{ji}^{(n)} + \sum_{n=0}^{N} {}_jp_{ii}^{(n)}.$$

Combining these four equations we obtain

$$\sum_{n=0}^{N} \{p_{ii}^{(n)} + p_{jj}^{(n)} - p_{ij}^{(n)} - p_{ji}^{(n)}\} = \sum_{\nu=1}^{N} f_{ij}^{(\nu)} \sum_{n=N-\nu+1}^{N} [p_{jj}^{(n)} - p_{ji}^{(n)}]$$
$$+ \sum_{\nu=N+1}^{\infty} f_{ij}^{(\nu)} \sum_{n=0}^{N} [p_{jj}^{(n)} - p_{ji}^{(n)}] + \sum_{n=0}^{N} {}_jp_{ii}^{(n)}.$$

As $N \to \infty$, the third term converges to $1 + {}_jp_{ii}^{*}$. The second term converges to zero by virtue of (6), since ${}_ip_{jj}^{*} < \infty$. For the first term we apply Proposition 6 with $a_\nu = f_{ij}^{(\nu)}$ and $b_n = \sum_{\nu=0}^{n} [p_{jj}^{(\nu)} - p_{ji}^{(\nu)}]$. The first condition in Proposition 6 is clearly satisfied, the second by virtue of (6). If i and j are in a null-recurrent class we have

$$\lim_{n\to\infty} (b_n - b_{n-1}) = \lim_{n\to\infty} (p_{jj}^{(n)} - p_{ji}^{(n)}) = 0 - 0 = 0.$$

If they are in an aperiodic, positive-recurrent class, the above limit is equal to

$$1/m_{jj} - 1/m_{ii} = 0$$

by Proposition 3. Hence in either case the third condition of Proposition 6 is also satisfied, and so the first term above converges to zero. Proposition 7 is proved.

When x_n is the sum of n independent and identically distributed integer-valued random variables with mean zero, Proposition 7 reduces to Theorem 1 in Spitzer [3]. His Theorem 2 can be proved in a similar way.

The following question is open, even in Spitzer's case: Is the series $\sum_n \{p_{ii}^{(n)} - p_{ji}^{(n)}\}$ convergent in a null-recurrent, equitable class? The corresponding problem for a positive-recurrent class has been considered in [1]. As a particular case in Theorem I. 11.4 there,[2] we have, if the class is aperiodic,

$$\sum_{n=0}^{\infty} \{p_{ii}^{(n)} - p_{ji}^{(n)}\} = m_{ji}/m_{ii}.$$

[2] I take this opportunity to correct a foolish slip in [1] about this theorem. If the class has period d, the proof there yields the convergence of $\sum_{\nu=1}^{nd+r} \{p_{jk}^{(\nu)} - p_{ik}^{(\nu)}\}$ as $n \to \infty$, for each fixed r. The theorem as stated is correct when $d = 1$; the corollary should be deleted.

If the class is also equitable, then it follows that the left member of (7) is equal to

$$m_{ji}/m_{ii} + m_{ij}/m_{jj} = (m_{ji} + m_{ij})/m_{ii}$$

by (3). The last-written fraction is well known to be equal to $1 + {}_jp_{ii}^*$, checking our previous result.

In a previous paper [2] I mentioned the problem of investigating the convergence of $\sum_n \{p_{ik}^{(n)} - p_{jk}^{(n)}\}$ in a null-recurrent class. Mr. C. J. Stone has recently informed me that this series need not converge. In this connection it may be worthwhile to record the following dual formulas. If $f_{ij}^* = 1$, then for each k,

$$\sum_{\nu=1}^{n} \{p_{ik}^{(\nu)} - p_{jk}^{(\nu)}\} = {}_jp_{ik}^* - \sum_{l \in I} {}_jp_{il}^* \, p_{lk}^{(n)};$$

if $e_{ij}^* = 1$, then for each k,

$$\sum_{\nu=1}^{n} \{p_{ki}^{(\nu)} - p_{kj}^{(\nu)}\} = \sum_{l \in I} p_{kl}^{(n)} \, {}_ip_{lj}^* - {}_ip_{kj}^* .$$

The first formula is in [1]; the second is proved dually.

Addendum. I am indebted to Hoeffding and Snell (independently) for the following extension of Proposition 7.

First, in formula (6), let i and j now belong to a recurrent class, but discard the assumption that $e_{ij}^* \leqq 1$, $e_{ji}^* \leqq 1$. Then without changing the proof there but noticing that $e_{ij}^* e_{ji}^* = 1$, we obtain in lieu of (6):

$$0 \leqq \sum_{n=0}^{N} [p_{jj}^{(n)} - e_{ij}^* \, p_{ji}^{(n)}] \leqq \sum_{n=0}^{N} {}_ip_{jj}^{(n)}, \qquad N \geqq 0. \tag{12}$$

Now let i and j be distinct states belonging to a null-recurrent class. Multiplying (8) and (11) through by e_{ij}^*, forming the following combination, and proceeding as before, we obtain in lieu of (7):

$$\sum_{n=0}^{\infty} \{e_{ij}^*(p_{ii}^{(n)} - p_{ji}^{(n)}) + p_{jj}^{(n)} - p_{ij}^{(n)}\} = e_{ij}^*(1 + {}_jp_{ii}^*) = 1 + {}_ip_{jj}^* . \tag{13}$$

Formula (13) is thus valid in any null-recurrent class and reduces to (7) when the class is also equitable.

REFERENCES

1. K. L. Chung, *Markov chains with stationary probabilities*, Berlin, Springer-Verlag, 1960.
2. ———, *Contributions to the theory of Markov chains*, J. Res. Nat. Bur. Standards, vol. 50 (1953), pp. 203–208.
3. F. Spitzer, *Some properties of recurrent random walk*, Illinois J. Math., vol. 5 (1961), pp. 234–245.

Syracuse University
Syracuse, New York

Reprinted with permission from *Proc. 4th Berkeley Sympos. Math. Stat. and Prob., II*, University of California Press, Berkeley, CA (1961), 35–56.

PROBABILISTIC METHODS IN MARKOV CHAINS

K. L. CHUNG
SYRACUSE UNIVERSITY

1. Introduction

To avoid constant repetition of qualifying phrases, we agree on the following notation, terminology, and conventions, unless otherwise specified.

I is a denumerable set of indices. The letters i, j, k, and l, with or without subscript, denote elements of **I**.

$\bar{\mathbf{I}} = \mathbf{I} \cup \{\infty\}$ is the one-point compactification of **I** considered as an isolated set of real numbers; $\infty > i$.

N is the set of nonnegative integers used as ordinals. The letters ν and n denote elements of **N**.

$\mathbf{T} = [0, \infty)$; $\mathbf{T}^0 = (0, \infty)$. The letters s, t and u, with or without subscript, denote elements of $\mathbf{T}^0$.

A statement or formula involving an unspecified element of **I** or $\mathbf{T}^0$ is meant to stand for every such element.

A sequence like $\{f_i\}$ is indexed by **I**; a matrix like (p_{ij}) is indexed by $\mathbf{I} \times \mathbf{I}$; a sum like $\sum_j$ is over **I**.

A function is real and finite valued. A function defined on $\mathbf{T}^0$ and having a right hand limit at zero is thereby extended to **T**; if in addition it is continuous in $\mathbf{T}^0$ it is said to be continuous in **T**.

A *(standard) transition matrix* is a matrix (p_{ij}) of functions on $\mathbf{T}^0$ satisfying the following conditions:

$$p_{ij}(t) \geqq 0, \tag{1.1}$$

$$\sum_j p_{ij}(t)p_{jk}(s) = p_{ik}(t + s), \tag{1.2}$$

$$\lim_{t \downarrow 0} p_{ii}(t) = 1, \tag{1.3}$$

$$\sum_j p_{ij}(t) = 1. \tag{1.4}$$

A *(temporally) homogeneous Markov chain*, or a *Markov chain with stationary transition probabilities*, associated with **I** and (p_{ij}), is a stochastic process $\{x_t\}$, $t \in \mathbf{T}$ or $t \in \mathbf{T}^0$, on the probability triple $(\Omega, \mathfrak{F}, \mathbf{P})$, with the generic sample point ω, having the following properties:

This paper was prepared with partial support of the Air Force Office of Scientific Research, under Contract AF 49(638)-265.

(1.5) *For each t in* $\mathbf{T}$ *or* $\mathbf{T}^0$ *respectively,* x_t *is a discrete random variable, and the set of all possible values of all* x_t *is* $\mathbf{I}$;

(1.6) *If* $t_1 < \cdots < t_n$, *then*

$$\mathbf{P}\{x(t_{\nu+1}, \omega) = i_{\nu+1}, 1 \leqq \nu \leqq n | x(t_1, \omega) = i_1\} = \prod_{\nu=1}^{n} p_{i_\nu i_{\nu+1}}(t_{\nu+1} - t_\nu).$$

An equivalent form of (1.6) is the *Markov property:*

$$(1.7) \qquad \mathbf{P}\{x(t_{n+1}, \omega) = i_{n+1} | x(t_\nu, \omega) = i_\nu, 1 \leqq \nu \leqq n\}$$
$$= \mathbf{P}\{x(t_{n+1}, \omega) = i_{n+1} | x(t_n, \omega) = i_n\} = p_{i_n i_{n+1}}(t_{n+1} - t_n).$$

A *version* of the process will be chosen to have the following further properties:

(1.8) *For any denumerable set R dense in* $\mathbf{T}$, *and every* $\omega \in \Omega$,

$$x(t, \omega) = \lim_{\substack{r \downarrow t \\ r \in R}} x(r, \omega)$$

for all t;

(1.9) *As a function of* (t, ω), $x(t, \omega)$ *is measurable with respect to the (uncompleted) product field* $\mathfrak{B} \times \mathfrak{F}$ *where* $\mathfrak{B}$ *is the usual Borel field on* $\mathbf{T}$.

The property (1.8) implies that the process is separable; the property (1.9) is called the Borel measurability of the process. Other properties of the process which follow from (1.5) to (1.9) for almost all ω, may be supposed to hold for all ω, so long as only denumerably many such properties are invoked.

From now on a process $\{x_t\}$ having the properties (1.5) to (1.9) will be abbreviated as an "M.C." It is called an *open* M.C. iff the parameter set is $\mathbf{T}^0$. The set $\mathbf{I}$ is called its *(minimal) state space*, the matrix (p_{ij}) its *transition matrix*. The distribution of x_0, when defined, is called its initial distribution $\{p_i\}$, where $p_i = \mathbf{P}\{\Delta_i\}$ and $\Delta_i = \{\omega : x(0, \omega) = i\}$. When $p_i = 1$, the resulting $\mathbf{P}$ will be written as $\mathbf{P}_i$; for example,

$$(1.10) \qquad \mathbf{P}_i\{x(t, \omega) = j\} = p_{ij}(t) = \mathbf{P}\{x(s + t, \omega) = j | x(s, \omega) = i\}$$

whenever the last is defined.

The study of the theory of M.C.'s consists in:

(a) uncovering the properties of, and relations among, the functions p_{ij};

(b) describing qualitatively and quantitatively the nature of the sample functions $x(\cdot, \omega)$, $\omega \in \Omega$; (less precisely, to analyze the evolution of the process in time).

Superficially at least, object (a) can be regarded as a purely "analytic" (as distinguished from "probabilistic" or "measure theoretic") program. We may simply wish to find as much information as possible about the set of functions satisfying (1.1) to (1.4). Or we may regard the matrices $\mathfrak{P}(t) = (p_{ij}(t))$ as forming a semigroup of operators and study the properties of the semigroup. A good number of papers have been written from such a standpoint eschewing probability itself "like the devil." For us however the most rewarding part of this

study is the interplay between the "analytic" and "stochastic" aspects of the theory. It is the main purpose of this paper to show, by various illustrations from recent work, that the structure of the transition matrix on the one hand, and the behavior of the sample functions on the other, are so intimately connected that one can hardly strike a chord in the one without bringing out an echo from the other. The two sides of the theory of Markov chains induce, sustain, and complement each other.

2. Comments on the conditions (1.3) and (1.4)

It has long been observed that much of the analytic structure of a transition matrix (p_{ij}) remains unchanged if the condition (1.4) is replaced by the weaker one

$$\sum_j p_{ij}(t) \leqq 1. \tag{2.1}$$

A matrix (p_{ij}) satisfying (1.1), (1.2), (1.3), and (2.1) will be called a *substochastic transition matrix*. (In distinction a transition matrix as defined in section 1 may be qualified as *stochastic*.) The above observation is easily justified by a simple reduction. Add a new index θ to $\mathbf{I}$ and define new elements as follows:

$$\begin{gathered} p_{i\theta}(t) = 1 - \sum_j p_{ij}(t), \\ p_{\theta\theta}(t) \equiv 1, \qquad p_{\theta j}(t) \equiv 0. \end{gathered} \tag{2.2}$$

The new matrix is stochastic and contains the old one. Probabilistically speaking, the new state θ is an absorbing state into which all the diminishing mass disappears. Thus $p_{i\theta}(t)$ is nondecreasing in t and we have

$$p_{i\theta}(t + s) - p_{i\theta}(t) = \sum_j p_{ij}(t)p_{j\theta}(s). \tag{2.3}$$

This trivial equation will assume more interesting proportions as we proceed.

Not only can the condition (1.4) be weakened into (2.1), but it can be dropped completely for many analytic purposes. This is implicit in some known proofs, but it was first realized in its full import by W. B. Jurkat [5] when he dispensed with this condition in more difficult cases. This realization has an important analytic consequence, for the omission of the "row condition" (1.4) restores complete symmetry to the rows and columns of the matrices. They form then simply a semigroup of nonnegative matrices $\{\mathfrak{P}(t)\}$ converging to the identity matrix I at $t = 0$. We shall not pursue the subject in this generality here since it has as yet no probabilistic interpretation.

Turning to the condition (1.3), let us first note that together with (1.1) and (1.2) it implies that every p_{ij} is continuous in $\mathbf{T}$ (see after lemma 1 below). Indeed, if we regard the semigroup $\{\mathfrak{P}(t)\}$ as operating on absolutely convergent series, then the condition (1.3) is equivalent to the strong continuity of the semigroup (see [4], p. 636). Now in the terminology of semigroup theory there is an even stronger kind of continuity, namely that in the "uniform oper-

ator topology," which is equivalent here to the condition that the convergence in (1.3) be uniform with respect to all $i \in \mathbf{I}$. Using the notation to be introduced at the beginning of section 5 below, it can be shown (theorem II. 19.2 of [1]) that this condition is equivalent to the boundedness of the sequence $\{q_i\}$. In this case the matrix $Q = (q_{ij})$ is a bounded operator and we have (see [4], p. 635)

$$\mathfrak{P}(t) = e^{Qt}, \tag{2.4}$$

Hence this case, which includes the case of a finite set $\mathbf{I}$, may be regarded as "solved" analytically. Probabilistically, the uniform condition implies (but is not implied by) that almost every sample function of the M.C. is a *step function*, namely one whose only discontinuities are jumps. While this was the case first studied for continuous parameter Markov processes, the properties of a sample step function are not essentially different from those of a sample sequence arising from a "discrete skeleton" (see section 6) of the M.C. The study of continuous parameter M.C.'s would scarcely be any innovation if we were to confine ourselves to this "trivial" case and label any new phenomenon as "pathological."

3. Two analytical lemmas

The first lemma is theorem II. 2.3 of [1], from which a superfluous condition has been removed, even though that very mild condition is satisfied in all known instances of application. The added argument is due to D. G. Austin (oral communication).

LEMMA 1. *Let (g_{ij}) be a matrix of nonnegative functions on $\mathbf{T}^0$ satisfying the condition that for every i,*

$$\lim_{t \downarrow 0} g_{ii}(t) = 1. \tag{3.1}$$

Let $\{f_j\}$ be nonnegative functions satisfying the following equations:

$$\begin{aligned} f_j(s+t) &= \sum_i f_i(s) g_{ij}(t), \qquad j \in \mathbf{I}, \\ [or \quad f_i(s+t) &= \sum g_{ij}(s) f_j(t), \qquad i \in \mathbf{I}]. \end{aligned} \tag{3.2}$$

Then each f_j is continuous in $\mathbf{T}$.

PROOF. It is proved in theorem II. 2.3 of [1] that each f_j is left-continuous and has a finite right-hand limit $f_j(t+0) \geqq f_j(t)$ for every $t \in \mathbf{T}^0$, and that $f_j(0+)$ exists. Such a function has at most a denumerable set D of discontinuities. If D is not empty, let $t_0 \in D$ so that $f_j(t_0 + 0) > f_j(t_0)$. Then there exist ϵ and δ_0 such that $f_j(t_0 + \delta) > f_j(t_0) + \epsilon$ if $0 < \delta < \delta_0$. There exists s_0 such that $g_{jj}(s) > 1/2$ if $0 < s < s_0$. Thus

$$f_j(t_0 + s + \delta) = \sum_i f_i(t_0 + \delta) g_{ij}(s) > [f_j(t_0) + \epsilon] g_{jj}(s) + \sum_{i \neq j} f_i(t_0 + \delta) g_{ij}(s). \tag{3.3}$$

Letting $\delta \downarrow 0$ and using Fatou's lemma, we have

$$(3.4)\qquad f_j(t_0+s+0) > \frac{\epsilon}{2} + \sum_i f_i(t_0)g_{ij}(s) = \frac{\epsilon}{2} + f_j(t_0+s).$$

Hence all points in (t_0, t_0+s_0) belong to D, a contradiction which proves the first part of the lemma. The second part is proved in the same way.

As a corollary we see that all p_{ij} satisfying (1.1), (1.2) and (1.3) are continuous in $\mathbf{T}$, without recourse to the condition (1.4). We remark however that with (1.4) or (2.1) each p_{ij} will be uniformly continuous in $\mathbf{T}$, which is not necessarily the case without it.

The second lemma is implicit in some previous work (see, for example, theorem II. 3.2 of [1]) but will be stated in a general form.

LEMMA 2. *Let (p_{ij}) be a matrix of functions satisfying* (1.1), (1.2), *and* (1.3). *Let $\{F_j\}$ be nonnegative, nondecreasing functions satisfying the equations*

$$(3.5)\qquad \begin{aligned} &F_i(s+t) - F_i(s) = \sum_j p_{ij}(s)F_j(t), \qquad i \in \mathbf{I},\\ [\textit{or}\quad &F_j(s+t) - F_j(t) = \sum_i F_i(s)p_{ij}(t), \qquad j \in \mathbf{I}]. \end{aligned}$$

Then each F_i has a continuous derivative F'_i satisfying

$$(3.6)\qquad \begin{aligned} &F'_i(s+t) = \sum_j p_{ij}(s)F'_j(t), \qquad t \in \mathbf{T}^0,\\ [\textit{or}\quad &F'_j(s+t) = \sum_i F'_i(s)p_{ij}(t), \qquad s \in \mathbf{T}^0]. \end{aligned}$$

REMARK. Taking the obvious differences, we see that the condition (3.5) is equivalent to the following: for any t_1 and t_2,

$$(3.7)\qquad F_i(s+t_2) - F_i(s+t_1) = \sum_j p_{ij}(s)[F_j(t_2) - F_j(t_1)].$$

PROOF. (A more elegant proof of this lemma has been given by Neveu [7].) By a theorem of Fubini on differentiation, we have for each s and almost all t,

$$(3.8)\qquad F'_i(s+t) = \sum_j p_{ij}(s)F'_j(t),$$

where F'_j denotes an almost everywhere derivative. Hence by Fubini's theorem on product measures, (3.8) is also true if $t \notin Z$ and $s \notin Z(t)$ where Z and $Z(t)$ are sets of Lebesgue measure zero. On the other hand we have by monotonicity and Fatou's lemma

$$(3.9)\qquad F'_i(s+t) \geqq \sum_j p_{ij}(s)F'_j(t)$$

for every s and t, if we agree now to take F'_j as the right-hand lower derivate. Let $t_0 \notin Z$ and suppose for a certain s_0 we have

$$(3.10)\qquad F'_i(s_0+t_0) > \sum_j p_{ij}(s_0)F'_j(t_0).$$

Then it follows that if $s > s_0$, since $p_{ii}(t) > 0$ for all t,

$$(3.11)\qquad \begin{aligned} F'_i(s+t_0) &\geqq \sum_j p_{ij}(s-s_0)F'_j(s_0+t_0) > \sum_j p_{ij}(s-s_0)\sum_k p_{jk}(s_0)F'_k(t_0)\\ &= \sum_k p_{ik}(s)F'_k(t_0). \end{aligned}$$

This is impossible by the second sentence of the proof; hence (3.8) must hold for all s, if $t_0 \notin Z$. For an arbitrary $t > 0$, let $t = t_0 + t_1$ where $t_0 \notin Z$. It follows that

$$F_i'(s+t) = F_i'(s+t_1+t_0) = \sum_j p_{ij}(s+t_1)F_j'(t_0) \tag{3.12}$$
$$= \sum_j \sum_k p_{ik}(s)p_{kj}(t_1)F_j'(t_0) = \sum_k p_{ik}(s)F_k'(t_1+t_0)$$
$$= \sum_k p_{ik}(s)F_k'(t).$$

Hence (3.8) holds for all $t > 0$, $s \geqq 0$. By the second part of lemma 1, each F_j' is continuous and consequently F_j has a continuous derivative. This proves the first part of lemma 2. The second part is proved in the same way.

4. Review of the strong Markov property

For a detailed discussion, see II. 8–9 of [1]. The reading of this section may be postponed until it becomes necessary.

Let $\{x_t\}$ be the M.C. defined in section 1. We denote by $\mathfrak{F}_t$ the augmented Borel field generated by $\{x_s, s \leqq t\}$. Let α be a nonnegative random variable with domain of definition Ω_α, where $\mathbf{P}(\Omega_\alpha) > 0$, which is "independent of the future," namely

$$\{\omega : \alpha(\omega) < t\} \in \mathfrak{F}_t \tag{4.1}$$

for every $t \in \mathbf{T}^0$. Such a random variable will be called *optional*. The Borel field of sets Λ (in $\mathfrak{F}$) such that for every t we have

$$\Lambda \cap \{\omega : \alpha(\omega) < t\} \in \mathfrak{F}_t \tag{4.2}$$

will be denoted by $\mathfrak{F}_\alpha$, the "past field relative to α." Let

$$y(t, \omega) = x[\alpha(\omega) + t, \omega], \qquad t \in \mathbf{T}^0. \tag{4.3}$$

It follows from (1.9) that $y_t = y(t, \cdot)$ with domain Ω_α is a random variable. The process $\{y_t, t \in \mathbf{T}^0\}$ will be called the *post-α process* and the augmented Borel field it generates will be denoted by $\mathfrak{F}_\alpha'$, "the future field relative to α." For any $\Lambda \in \mathfrak{F}_\alpha$ we put

$$A(\Lambda; t) = \mathbf{P}\{\Lambda; \alpha(\omega) \leqq t\}. \tag{4.4}$$

The measure corresponding to this distribution function will be called the $A(\Lambda; \cdot)$ measure.

The following collection of assertions, valid for each optional α, will be referred to as the *strong Markov property*.

(1) For every $\Lambda \in \mathfrak{F}_\alpha$ and $\mathrm{M} \in \mathfrak{F}_\alpha'$ we have

$$\mathbf{P}\{\Lambda\mathrm{M}|y_0\} = \mathbf{P}\{\Lambda|y_0\}\mathbf{P}\{\mathrm{M}|y_0\} \tag{4.5}$$

almost everywhere on the set $\{\omega : y_0(\omega) \in \mathbf{I}\}$.

(2) The post-α process $\{y_t, t \in \mathbf{T}^0\}$ is an open M.C. which has the properties

corresponding to (1.8) and (1.9), and whose transition matrix is a part of (p_{ij}). In particular, $\{y_t, t \in \mathbf{T}\}$ is a M.C. on the set $\{\omega : y_0(\omega) \in \mathbf{I}\}$.

(3) For each $j \in \mathbf{I}$, $\Lambda \in \mathfrak{F}_\alpha$ and almost every $s \in (0, t)$ with respect to the $A(\Lambda; \cdot)$ measure, we have

$$\mathbf{P}\{x(t, \omega) = j|\Lambda; \alpha(\omega) = s\} = \mathbf{P}\{y(t - s, \omega) = j|\Lambda; \alpha(\omega) = s\}. \tag{4.6}$$

One version of the conditional probability in (4.6), to be denoted by $r_j(s, t|\Lambda)$, is continuous in $t \in [s, \infty)$ for each $s \in \mathbf{T}$.

The following particular case of the strong Markov property, to be referred to as the *strongest Markov property*, will be applied in the sequel. The two fields $\mathfrak{F}_\alpha$ and $\mathfrak{F}'_\alpha$ are said to be independent iff for every $\Lambda \in \mathfrak{F}_\alpha$ and $M \in \mathfrak{F}'_\alpha$ we have

$$\mathbf{P}\{\Lambda M|\Omega_\alpha\} = \mathbf{P}\{\Lambda|\Omega_\alpha\}\mathbf{P}\{M|\Omega_\alpha\}; \tag{4.7}$$

alternately, since $\Lambda \in \Omega_\alpha$,

$$\mathbf{P}\{\Lambda M\} = \mathbf{P}\{\Lambda\}\mathbf{P}\{M|\Omega_\alpha\}. \tag{4.8}$$

(4) The fields $\mathfrak{F}_\alpha$ and $\mathfrak{F}'_\alpha$ are independent if and only if there exist functions $\{\rho_j\}$ on $\mathbf{T}^0$ such that for every $j \in \mathbf{I}$, $t \in \mathbf{T}^0$ and $\Lambda \in \mathfrak{F}_\alpha$ we have

$$r_j(s, t|\Lambda) = \rho_j(t - s) \tag{4.9}$$

for almost all s in $(0, t)$ with respect to the $A(\Lambda; \cdot)$ measure. We have then

$$\rho_j(t) = \mathbf{P}\{y(t, \omega) = j|\Omega_\alpha\} \tag{4.10}$$

and ρ_j is continuous in $\mathbf{T}$.

In particular, this is the case if for a fixed j we have

$$\mathbf{P}\{y(0, \omega) = j|\Omega_\alpha\} = 1. \tag{4.11}$$

5. Transition from and to a stable state

Let us introduce the following notation:

$$-p'_{ii}(0) = \lim_{t \downarrow 0} \frac{1 - p_{ii}(t)}{t} = -q_{ii} = q_i \leqq \infty, \tag{5.1}$$

$$p'_{ij}(0) = \lim_{t \downarrow 0} \frac{p_{ij}(t)}{t} = q_{ij} < \infty, \qquad i \neq j. \tag{5.2}$$

That these limits exist and have the indicated finiteness is well known (theorems II. 2.4 and II. 2.5 of [1]). Analytically, (5.1) follows from the subadditivity of $-\log p_{ii}(t)$ which is a consequence of (1.1), (1.2), and (1.3) without the intervention of (1.4). The corresponding basic property of sample functions is given in the formula

$$\mathbf{P}_i\{x(s, \omega) \equiv i, 0 < s < t\} = e^{-q_i t}, \tag{5.3}$$

where the right member stands for 0 if $q_i = \infty$ and $t > 0$. The state i is called *stable* or *instantaneous* according as $q_i < \infty$ or $q_i = \infty$.

In the rest of this section let i be fixed and $p_i = 1$ so that $\mathbf{P} = \mathbf{P}_i$. Define on Δ_i the *first exit time from i:*

$$\alpha(\omega) = \inf \{t : t > 0,\ x(t, \omega) \neq i\}. \tag{5.4}$$

Then (5.3) is equivalent to the assertion that α is a random variable with the distribution function

$$e_{q_i}(t) \overset{\text{def}}{=} 1 - e^{-q_i t}, \qquad t \in \mathbf{T}^0, \tag{5.5}$$

which reduces to the unit distribution ϵ if $q_i = \infty$. It is easy to see that α is optional. It may or may not be easy to see that $\mathfrak{F}_\alpha$ and $\mathfrak{F}'_\alpha$ are independent in the sense of (4) of section 4. For a tedious but rigorous proof of this fact, see theorem II. 15.2 of [1]; a partially analytic proof will be given later.

We have as a trivial identity valid for any α:

$$\mathbf{P}\{x(t, \omega) = j\} = \mathbf{P}\{\alpha(\omega) \leqq t; x(t, \omega) = j\} + \mathbf{P}\{\alpha(\omega) > t; x(t, \omega) = j\}. \tag{5.6}$$

Now let i be a stable state. The second term above is $\delta_{ij} \exp(-q_i t)$ by (5.3). The first term may be written as

$$\int_0^t \mathbf{P}\{x(t, \omega) = j | \alpha(\omega) = s\}\, d\mathbf{P}\, \{\alpha(\omega) \leqq s\} \tag{5.7}$$

by the definition of conditional probability. By (4) of section 4, and writing r_{ij} for the ρ_j then we see that (5.6) becomes

$$p_{ij}(t) = \int_0^t r_{ij}(t - s) q_i e^{-q_i s}\, ds + e^{-q_i t}\delta_{ij}. \tag{5.8}$$

Furthermore by (2) of section 4, we have

$$r_{ik}(t + s) = \sum_j r_{ij}(t) p_{jk}(s), \tag{5.9}$$

$$\sum_j r_{ij}(t) = 1. \qquad k \in \mathbf{I};\ s, t \in \mathbf{T}^0. \tag{5.10}$$

The above formulas give an integral representation of p_{ij} obtained by a precise analysis of the local behavior of a sample function at the exit from the stable state i. It is a clear example of the probabilistic method in reaching analytic conclusions.

For it follows from (5.9) and lemma 1 that r_{ij} is continuous in $\mathbf{T}$. It is then an immediate consequence of (5.8) that p_{ij} has a continuous derivative p'_{ij} satisfying the following:

$$e^{-q_i t} \frac{d}{dt} [e^{q_i t} p_{ij}(t)] = p'_{ij}(t) + q_i p_{ij}(t) = q_i r_{ij}(t). \tag{5.11}$$

It follows furthermore from (1.2), (1.4), (5.9), and (5.10) that

$$\sum_j p'_{ij}(t) = 0, \tag{5.12}$$

$$\sum_j |p'_{ij}(t)| \leqq 2q_i, \qquad t \in \mathbf{T}^0, \tag{5.13}$$

$$\sum_k p'_{ik}(t) p_{kj}(s) = p'_{ij}(t + s), \tag{5.14}$$

namely that both the series in (1.2) and (1.4) can be differentiated term by term in T^0 to yield absolutely convergent series—a by no means trivial analytical fact. Our proof shows that this is tied up with the fact that the post-α process is Markovian with the same transition matrix (curtailed). The critical case for $t = 0$ will be examined later in section 7.

The formula (5.8) has a dual which will be briefly discussed. Let j be stable and i arbitrary, then we have

$$p_{ij}(t) = \delta_{ij}e^{-q_it} + \int_0^t v_{ij}(s)e^{-q_j(t-s)}\,ds. \tag{5.15}$$

The function v_{ij} represents a renewal density function; precisely v_{ij} is the derivative of V_{ij} where $V_{ij}(t)$ is the expected number of entrances into the state j in the open interval $(0, t)$, under the hypothesis $p_i = 1$. Using the notation of section 6, we have in fact

$$V_{ij}(t) = \sum_{n=0}^{\infty} [F_{ij} * F_{jj}^{n*}](t), \tag{5.16}$$

where $*$ denotes the convolution of distribution functions, $F_{jj}^{0*} = \epsilon$, and $F_{jj}^{(n+1)*} = F_{jj}^{n*} * F_{jj}$; but this explicit formula will not be needed. From the probabilistic meaning we infer that

$$V_{ij}(s + t) - V_{ij}(s) = \sum_k p_{ik}(s)V_{kj}(t). \tag{5.17}$$

The existence of the continuous derivative v_{ij} follows from (5.17) and lemma 2. (This is a better approach than that in section II. 16 of [1].) Furthermore it follows from (5.14) that

$$\frac{d}{dt}[p_{ij}(t)e^{q_jt}]e^{-q_jt} = p'_{ij}(t) + p_{ij}(t)q_j = v_{ij}(t); \tag{5.18}$$

$$\sum_k p_{ik}(s)p'_{kj}(t) = p'_{ij}(s + t), \tag{5.19}$$

where the series converges absolutely.

Having deduced the preceding results by probabilistic methods, we are now ready for an analytic short cut based on hindsight. The fact that $(\exp q_it)p_{ij}(t)$ is nondecreasing in t, as shown in (5.11), can be proved directly as follows. Since $p_{ii}(h) \geq \exp(-q_ih)$ by the subadditivity mentioned in connection with (5.1) [or probabilistically as a consequence of (5.3)], we have

$$e^{q_i(t+h)}p_{ij}(t + h) \geq e^{q_ih}p_{ii}(h)e^{q_it}p_{ij}(t) \geq e^{q_it}p_{ij}(t). \tag{5.20}$$

Let $P_{ij}(t) = \int_0^t p_{ij}(s)\,ds$. Then we have by partial integration,

$$p_{ij}(t) - \delta_{ij} + q_iP_{ij}(t) = \int_0^t e^{-q_is}D[e^{q_is}p_{ij}(s)]\,ds, \tag{5.21}$$

where D denotes an almost everywhere derivative. Since this derivative is nonnegative, the left member of (5.21) is a nondecreasing function of t. Now a trivial calculation based on (1.2) yields

$$(5.22)\qquad \sum_j [p_{ij}(t) - \delta_{ij} + q_i P_{ij}(t)]p_{jk}(s) = p_{ik}(t+s) + q_i P_{ik}(t+s) - p_{ik}(s) - q_i P_{ik}(s).$$

Thus the conditions for the second part of lemma 2 are satisfied if we take $F_j(t)$ to be the left member of (5.21). It follows that p_{ij} has a continuous derivative satisfying (5.14). In an exactly dual way (5.19) can be proved. We remark also that neither proof utilizes (1.4).

As far as the analytic part is concerned, the above approach is the simplest. We can now retrace our steps to *define* r_{ij} by means of the second equation in (5.11), verify (5.8) as a consequence, and using (4) of section 4, conclude that the two fields $\mathfrak{F}_\alpha$ and $\mathfrak{F}'_\alpha$ are independent.

We add the following remarks before turning to another illustration of this kind. The rather complete success of the methods developed in this section depends on the primary fact that the *set of constancy*,

$$(5.23)\qquad S_i(\omega) = \{t : x(t, \omega) = i\}$$

for a fixed stable i, consists of a sequence of disjoint intervals without clustering in the finite (theorem II. 5.7 of [1]). Thus the endpoints of these intervals form natural relay points in the analysis of the sample functions, with the length of an interval (sojourn time) corresponding analytically to the smoothing exponential factor $\exp(\pm q_i t)$. It is not known whether suitable substitutes for (5.11) and (5.18), or (5.8) and (5.15), exist in the general case where both i and j are arbitrary. On the other hand, it has been proved by D. Ornstein [8] (see also Jurkat [5] and the appendix in [1]) that the equations (5.12), (5.14), and (5.19) remain valid in the general case. This can be proved by the development in the next section.

6. First entrance and last exit

Let $i \neq j$ and let Δ_{ij} be the subset of Δ_i where the following infimum is finite:

$$(6.1)\qquad \alpha_{ij}(\omega) = \inf\{t : t > 0, x(t, \omega) = j\}.$$

It is verified that α_{ij} is an optional random variable, and in view of the last sentence in section 4, the strongest Markov property applies with $y(0, \omega) = j$ on Δ_{ij}, and the ρ_j in (4.10) reducing to p_{jj} (in general $\rho_k = p_{jk}$). Now if $\alpha = \alpha_{ij}$ in (5.6), the second term vanishes by definition and we obtain, by what has just been said,

$$(6.2)\qquad p_{ij}(t) = \int_0^t p_{jj}(t-s)\, dF_{ij}(s),$$

where

$$(6.3)\qquad F_{ij}(t) = \mathbf{P}_i\{\alpha_{ij}(\omega) \leq t\}.$$

It is easy to see that F_{ij} is continuous in $\mathbf{T}$ but more will be shown presently. The formula (6.2) is the *first entrance formula from i to j*. The definitions (6.1) and (6.3) may be extended to the case $i = j$, yielding $F_{ii}(t) \equiv 1$. The last definition,

as well as (6.4) below, differs from that given in section II.11 of [1] but the latter agrees with that in the appendix there.

To proceed further we must introduce the *taboo probability functions*

$$(6.4) \qquad {}_jp_{ik}(t) = \mathbf{P}_i\{x(t, \omega) = k;\, x(s, \omega) \neq j,\, 0 < s < t\}.$$

It follows from the stochastic continuity of the M.C. [equivalent to condition (1.3)] that ${}_jp_{ik}(t) \equiv 0$ if $i = j$ or $k = j$. These probabilities are well defined on account of the separability of the process. We observe that

$$(6.5) \qquad F_{ij}(t) = 1 - \sum_k {}_jp_{ik}(t), \qquad i \neq j.$$

For fixed j, the matrix $({}_jp_{ik})$ with i and k in $\mathbf{I} - \{j\}$, is a substochastic transition matrix:

$$(6.6) \qquad \sum_k {}_jp_{ik}(t)\,{}_jp_{kl}(s) = {}_jp_{il}(t + s).$$

It is unnecessary to exclude j from the summation since the corresponding term vanishes. For this substochastic transition matrix, F_{ij} plays the role of $p_{i\theta}$ in section 2. It follows at once [compare (2.3)] that

$$(6.7) \qquad F_{ij}(s + t) - F_{ij}(s) = \sum_k {}_jp_{ik}(s)F_{kj}(t), \qquad i \neq j.$$

Hence an application of lemma 2 shows that each F_{ij} has a continuous derivative f_{ij} satisfying

$$(6.8) \qquad f_{ij}(s + t) = \sum_k {}_jp_{ik}(s)f_{kj}(t),$$

and consequently (6.2) can be improved into

$$(6.9) \qquad p_{ij}(t) = \int_0^t f_{ij}(s)p_{jj}(t - s)\, ds, \qquad i \neq j.$$

It turns out that the formula (6.9) has a dual which has been proved in general only recently (the case where i is stable being previously known). To motivate this dualization it is best to consider the discrete parameter analogues.

For each $h \in \mathbf{T}^0$ the stochastic process $\{x_{nh}, n \in \mathbf{N}\}$ is called the ***discrete skeleton*** of $\{x_t, t \in \mathbf{T}\}$ at the scale h. It is a discrete parameter homogeneous Markov chain with the *n-step transition matrix* $(p_{ij}^{(n)})$. Let

$$(6.10) \qquad {}_jp_{ik}^{(n)}(h) = \mathbf{P}_i\{x(nh, \omega) = k,\, x(\nu h, \omega) \neq j,\, 1 \leqq \nu \leqq n - 1\}$$

be the corresponding taboo probabilities. The analogue of (6.9) is then

$$(6.11) \qquad p_{ij}^{(n)}(h) = \sum_{\nu=1}^{n} {}_jp_{ij}^{(\nu)}(h)p_{jj}^{(n-\nu)}(h), \qquad n \geqq 1,$$

where ${}_jp_{ij}^{(\nu)}(h)$ may be denoted by $f_{ij}^{(\nu)}(h)$ for comparison with (6.9) but is preferably written as shown with a view to dualization. This is a very old formula and is basic in the so-called theory of "recurrent events" (see section I. 8 of [1]). Now in the discrete parameter case the reasoning leading to (6.11) can be immediately dualized by interchanging "i" and "j," "first" and "last," "entrance" and "exit," to yield the dual:

$$p_{ij}^{(n)}(h) = \sum_{\nu=0}^{n-1} p_{ii}^{(\nu)}(h)\,{}_ip_{ij}^{(n-\nu)}(h), \qquad n \geq 1. \tag{6.12}$$

These two formulas (6.11) and (6.12), valid also for $i = j$, are particular cases of theorem I. 9.1 of [1]. Since the taboo probabilities can be defined algebraically, they appear as simple algebraic consequences of the operation of matrix multiplication, apart from questions of convergence. Now if (1.4) or the weaker (2.1) holds, then

$$\sum_{n=1}^{\infty} {}_jp_{ij}^{(n)}(h) \leq 1, \tag{6.13}$$

which greatly facilitates the passage to limit in (6.11) as $h \downarrow 0$. The same however cannot be said of the series $\sum_{n=1}^{\infty} {}_ip_{ij}^{(n)}(h)$. Thus it is desirable to execute the limit operation without the advantage of (1.4), but making defter use of (1.3). The main idea is to consider a *sequence* of $h \downarrow 0$ such that

$$\sum_{nh \leq t} {}_jp_{ij}^{(n)}(h) \quad \text{and} \quad \sum_{nh \leq t} {}_ip_{ij}^{(n)}(h) \tag{6.14}$$

converge for a dense set of t, in the manner of Helly's selection principle. This is carried out by Jurkat [5] with a further refinement.

While this method has analytic power, it is unfortunately devoid of probabilistic meaning at the moment. We shall sketch two different approaches based on considerations of sample functions.

Since (6.9) is obtained by analyzing the first entrance into the final state j, it is natural to reflect upon the last exit from the initial state i. Let us define on Δ_i:

$$\gamma_i(t, \omega) = \sup \{s : 0 \leq s \leq t,\ x(s, \omega) = i\}. \tag{6.15}$$

For each fixed t this is a random variable but clearly it is not optional in any sensible way: to determine if $\gamma_i(t, \omega) \leq s$ we must know $x(\cdot, \omega)$ up to the time t. On the other hand, its distribution function is easily written down, if $0 \leq s < t$,

$$\Gamma_i(s, t) \overset{\text{def}}{=} \mathbf{P}_i\{\gamma_i(t, \omega) \leq s\} = \sum_k p_{ik}(s)[1 - F_{ki}(t - s)]. \tag{6.16}$$

Furthermore, for every $j \neq i$ we have

$$\Gamma_{ij}(s, t) \overset{\text{def}}{=} \mathbf{P}_i\{\gamma_i(t, \omega) \leq s;\ x(t, \omega) = j\} = \sum_k p_{ik}(s)\,{}_ip_{kj}(t - s) \tag{6.17}$$

so that, for $0 \leq s < t$, we have

$$\Gamma_i(s, t) = \sum_{j \neq i} \Gamma_{ij}(s, t). \tag{6.18}$$

For $s = t$ the above equation becomes false. We have

$$p_{ij}(t) = \int_0^t \mathbf{P}_i\{x(t, \omega) = j | \gamma_i(t, \omega) = s\}\, d_s\, \Gamma_i(s, t). \tag{6.19}$$

Now the salient fact here is that the conditional probability in (6.19) turns out to be a function of $t - s$ only, while the distribution function $\Gamma_i(s, t)$ has a density function which is the product of a function of $t - s$ and one of s only.

To demonstrate these facts by our first method, we decompose the sample functions $x(\cdot, \omega)$ with $x(0, \omega) = i$ and $x(t, \omega) = j$ into subsets according to the location of $\gamma_i(t, \omega)$. To be precise, for each n let $\gamma_i^{(n)}(t, \omega)$ be the unique dyadic number $(\nu - 1)2^{-n}$ such that

$$x[(\nu - 1)2^{-n}, \omega] = i \quad \text{and} \quad x(u, \omega) \neq i, \qquad \nu 2^{-n} \leqq u \leqq t. \tag{6.20}$$

We have $\lim_{n\to\infty} \gamma_i^{(n)}(t, \omega) = \gamma_i(t, \omega)$ by separability, and consequently

$$\begin{aligned} p_{ij}(t) &= \lim_{n\to\infty} \sum_{\nu \leqq 2^n t} \mathbf{P}_i\{\gamma_i^{(n)}(t, \omega) = (\nu - 1)2^{-n}; x(t, \omega) = j\} \\ &= \lim_{n\to\infty} \sum_{\nu \leqq 2^n t} p_{ii}((\nu - 1)2^{-n}) \sum_k p_{ik}(2^{-n})\,{}_ip_{kj}(t - \nu 2^{-n}). \end{aligned} \tag{6.21}$$

The last written sum may be exhibited as

$$\int_0^t \phi_{ij}^{(n)}(t - s)\, d\, \pi_i^{(n)}(s) \tag{6.22}$$

where

$$\pi_i^{(n)}(s) = \sum_{\nu \leqq 2^n s} p_{ii}((\nu - 1)2^{-n})2^{-n}, \qquad \phi_{ij}^{(n)}(s) = 2^n \sum_k p_{ik}(2^{-n})\,{}_ip_{kj}(s). \tag{6.23}$$

Clearly,

$$\lim_{n\to\infty} \pi_i^{(n)}(t) = \int_0^t p_{ii}(s)\, ds. \tag{6.24}$$

Hence it remains to show that $\phi_{ij}^{(n)}(s)$ converges *uniformly* in every finite interval to $g_{ij}(s)$ in order to obtain in the limit the desired formula:

$$p_{ij}(t) = \int_0^t g_{ij}(t - s)p_{ii}(s)\, ds. \tag{6.25}$$

By the definition of $\phi_{ij}^{(n)}(s)$,

$$\sum_j g_{ij}(s)\,{}_ip_{jk}(t) = g_{ik}(s + t), \tag{6.26}$$

and so by lemma 1 all q_{ij} are continuous in $\mathbf{T}$. The convergence of $\phi_{ij}^{(n)}$ follows from properties of taboo probability functions, only the uniformity causes some technical difficulty. This plan of attack has been carried out in detail in [1]. The purpose of the résumé above is to show the basic probabilistic idea underlying this method.

Our second method shows promise of general applicability, being inherent in the nature of the stochastic scheme of things. It is that of reversing the direction of time, or retracing the process. Formally let $U \in \mathbf{T}^0$ and define

$$z^U(t, \omega) = x(U - t, \omega), \qquad 0 \leqq t \leqq U. \tag{6.27}$$

The new process $\{z_t^U, 0 \leqq t \leqq U\}$ is Markovian with the state space $\mathbf{I}$, but has in general *nonstationary* transition probabilities. This is one difficulty to be faced in this approach, the other one being the dependence on U. But these difficulties may also give us new clues.

For the sake of simplicity let us suppose that $p_i = 1$. Then if $0 \leqq s \leqq t \leqq U$, we have

$$p^U(s, t; j, i) \overset{\text{def}}{=} \mathbf{P}\{z^U(t, \omega) = i | z^U(s, \omega) = j\} = \frac{p_{ii}(U - t)}{p_{ij}(U - s)} p_{ij}(t - s). \tag{6.28}$$

The first entrance time distribution form j to i, starting at time s, is also easily written down:

$$\begin{aligned} F^U(s, t; j, i) &\overset{\text{def}}{=} \mathbf{P}\{z^U(u, \omega) = i \quad \text{for some} \quad u \in [s, t] | z^U(s, \omega) = j\} \\ &= 1 - \frac{1}{p_{ij}(U - s)} \sum_k p_{ik}(U - t)_i p_{kj}(t - s) \\ &= 1 - \frac{1}{p_{ij}(U - s)} \Gamma_{ij}(U - t, U - s). \end{aligned} \tag{6.29}$$

Now the reversed Markov chain (if the proper version is taken) also possesses a strong Markov property, a particular case of which is the first entrance formula generalizing (6.2),

$$p^U(s, t; j, i) = \int_s^t p^U(u, t; i, i)\, d_u F^U(s, u; j, i). \tag{6.30}$$

For a proof of this see [2]. Substituting from (6.28) and (6.29) we obtain

$$p_{ij}(t - s) = \int_t p_{ii}(t - u) \frac{d_u \Gamma_{ij}(U - u, U - s)}{p_{ii}(U - u)}, \tag{6.31}$$

or

$$p_{ij}(t) = \int_t^0 p_{ii}(t - u) \frac{d_u \Gamma_{ij}(U - s - u, U - s)}{p_{ii}(U - s - u)}, \tag{6.32}$$

if $t \leqq U - s$. This being so it is reasonable to conjecture that the measures in u generated by $\Gamma_{ij}(U - u, U)/p_{ii}(U - u)$ for different values of $U - u$ coincide, namely, there exists a nondecreasing function G_{ij} on $\mathbf{T}$ such that

$$\int_{u_2}^{u_1} \frac{d_u \Gamma_{ij}(U - u, U)}{p_{ii}(U - u)} = \int_{u_1}^{u_2} d\, G_{ij}(u) \tag{6.33}$$

for $0 \leqq u_1 \leqq u_2 \leqq U$. This is indeed true by a known, though formidable, theorem due to Titchmarsh [10], p. 328. (For a proof by real variable method, see Mikusinski [6], chapter 7.) We have by (6.17) and (6.6)

$$\sum_j \Gamma_{ij}(U - u, U)_i p_{jk}(s) = \Gamma_{ik}(U - u, U + s). \tag{6.34}$$

Hence for $0 \leqq u_2 \leqq u_1 \leqq U - s$,

$$\begin{aligned} \int_{u_2}^{u_1} \sum_j \frac{d_u \Gamma_{ij}(U - u, U)}{p_{ii}(U - u)}\, {}_i p_{jk}(s) &= \int_{u_2}^{u_1} \frac{d_u \Gamma_{ik}(U - u, U + s)}{p_{ii}(U - u)} \\ &= \int_{s+u_2}^{s+u_1} \frac{d_u \Gamma_{ik}(U + s - u, U + s)}{p_{ii}(U + s - u)} \\ &= \int_{s+u_2}^{s+u_1} \frac{d_u \Gamma_{ik}(U - u, U)}{p_{ii}(U - u)}, \end{aligned} \tag{6.35}$$

where the last equation follows from (6.33). This is equivalent to

$$(6.36)\qquad \sum_j [G_{ij}(u_2) - G_{ij}(u_1)]_i p_{jk}(s) = G_{ik}(s + u_2) - G_{ik}(s + u_1).$$

Thus by lemma 2 (see the remark there) each G_{ij} has a continuous derivative g_{ij} in $\mathbf{T}$ satisfying (6.26). Substituting back into (6.32) we obtain (6.25).

Incidentally, we have shown that $\Gamma_{ij}(U - u, U)$ has a derivative with respect to u in $[0, U]$ which is equal to $p_{ii}(U - u)g_{ij}(u)$, verifying the remark after (6.19). This can also be deduced from (6.25) and (6.26) since by (6.17) we have

$$(6.37)\qquad \Gamma_{ij}(s, t) = \sum_k \int_0^s p_{ii}(u)g_{ik}(s - u)_i p_{kj}(t - s)\, du$$

$$= \int_0^s p_{ii}(u)g_{ij}(t - u)\, du.$$

Summing (6.25) over all $j \neq i$, we see that

$$(6.38)\qquad 1 - p_{ii}(t) = \int_0^t g_i(t - s)p_{ii}(s)\, ds,$$

where $g_i = \sum_{j\neq i} g_{ij}$. This integral equation for p_{ii} can be made as the starting point of another proof of Ornstein's theorem [8] on the continuous differentiability in $\mathbf{T}^0$ of all p_{ij} of a transition matrix. Such a proof is given by Jurkat [5] without the use of (1.4). He has also indicated a proof which is based on (6.9) instead of (6.25). It can be shown moreover that the series in (5.14) and (5.19) converge absolutely in $\mathbf{T}^0$ [without the condition (1.4)] and so does that in (5.12) if (1.4) is assumed. However the following problem is open: if i is instantaneous, is it true that

$$(6.39)\qquad \lim_{t \downarrow 0} p'_{ii}(t) = -q_i = -\infty ?$$

The answer is "yes" if i is stable, as a consequence of (5.11) and the existence of $r_{ii}(0+) = 0$. This problem is particularly interesting since almost every sample function $x(\cdot, \omega)$ with $x(0, \omega) = i$ "oscillates tremendously" at $t = 0$, while it is not even known if p_{ii} is monotone in a neighborhood of zero.

I take this opportunity to correct an oversight (p. 270 of [1], lines 4 to 5) brought to my attention by Reuter. For every i and j, we have

$$(6.40)\qquad \lim_{t\to\infty} p'_{ij}(t) = 0.$$

This follows by fixing a positive t in equation (27) there, let $s \to \infty$ according to theorem II. 10.1, and use the inequality in (28) to justify uniform convergence with respect to s. The existence of the limit in (6.38) implies that it is equal to zero.

7. The minimal chain

Returning to section 5, we now wish to study what happens at the exact moment of exit from a stable state i. Noting that (4.10) remains in force at $t = 0$ but, instead of (5.10), we have by Fatou's lemma

(7.1) $$\mathbf{P}\{y(0, \omega) \in \mathbf{I}|\Omega_\alpha\} = \sum_j r_{ij}(0) \leqq 1.$$

Since

(7.2) $$r_{ij}(0) = \frac{(1 - \delta_{ij})q_{ij}}{q_i}$$

by (5.2) and (5.11), this amounts to the easy analytic result

(7.3) $$\sum_{j \neq i} q_{ij} \leqq q_i.$$

If strict inequality holds above, then with a probability equal to $1 - \sum_j r_{ij}(0) > 0$ we have $y(0, \omega) = \infty$. We recall that ∞ is the "point at infinity" adjoined to compactify $\mathbf{I}$ to render the process separable. For a general optional α and the post-α process $\{y_t\}$, $y(0, \omega) = \infty$ if and only if $\lim_{t \downarrow \alpha(\omega)} x(t, \omega) = \infty$, on account of (1.8). On the set of ω for which this is true the process $\{y_t\}$ does not have an initial distribution (on $\mathbf{I}$), and is a Markov chain only in $\mathbf{T}^0$; see (2) of section 4. It is important to note that

(7.4) $$\mathbf{P}\{y(t, \omega) \in \mathbf{I}|\Omega_\alpha\} = 1, \qquad t \in \mathbf{T}^0,$$

is part of the assertion of the strong Markov property. The above conclusions may be stated as follows: at the first exit time $\alpha(\omega)$ from the stable state i, the probability of a *pseudojump to* j ($\neq i$) is $r_{ij}(0) = q_{ij}/q_i$, and the probability of a *pseudojump to* ∞ is $1 - \sum_{j \neq i} (q_{ij}/q_i)$. We say "pseudojump" rather than "jump," since if j is instantaneous the sample function does not have a jump in the usual sense but shows the following behavior,

(7.5) $$\varliminf_{t \uparrow \alpha(\omega)} x(t, \omega) = j < \infty = \varlimsup_{t \downarrow \alpha(\omega)} x(t, \omega).$$

We have thus a complete analysis of the first discontinuity of a sample function which starts at a stable state. To continue this process, we shall assume that all states are stable and that equality holds in (7.1) or (7.3) so that a pseudojump to j is a genuine jump and the possibility of a pseudojump to ∞ is excluded. Finally we suppose that there is no *absorbing state* to omit trivial modifications. These assumptions are summed up as follows:

(7.6) $$0 < q_i = \sum_{j \neq i} q_{ij} < \infty, \qquad i \in \mathbf{I}.$$

The preceding analysis then implies, by an induction on the number of jumps, that there are infinitely many jumps of the sample function

(7.7) $$\tau_1(\omega) < \cdots < \tau_n(\omega) < \cdots.$$

Let us put also $\tau_0(\omega) = 0$ and

(7.8) $$\chi_n(\omega) = x(\tau_n(\omega), \omega) = \lim_{t \downarrow \tau_n(\omega)} x(t, \omega).$$

It is easy to verify that each τ_n is optional with $\mathbf{P}\{\Omega_{\tau_n}\} = 1$. (One may use in this connection theorem II. 15.1 of [1], but that is not necessary.) It follows from (1) of section 4 that

$$(7.9)\qquad \mathbf{P}\{\chi_{n+1}(\omega) = i_{n+1}|\chi_\nu(\omega) = i_\nu, 0 \leqq \nu \leqq n\} = \mathbf{P}\{\chi_{n+1}(\omega) = i_{n+1}|\chi_n(\omega) = i_n\}.$$

Applying the preceding analysis of the first discontinuity to the post-τ_n process, we see that the right member of (7.9) is equal to $r_{i_n i_{n+1}}(0)$. Hence $\{\chi_n, n \in N\}$ is a discrete parameter homogeneous Markov chain with the one-step transition matrix $(r_{ij}(0))$. Furthermore it follows from (5.3) applied to the post-τ_n process that

$$(7.10)\qquad \mathbf{P}\{\tau_{n+1}(\omega) - \tau_n(\omega) \leqq t|\chi_\nu(\omega), 0 \leqq \nu \leqq n-1; \chi_n(\omega) = i\} = e_{q_i}(t).$$

Now let

$$(7.11)\qquad \tau_\infty(\omega) = \lim_{n\to\infty} \tau_n(\omega).$$

Then it is clear from the definition that for almost all ω,

$$(7.12)\qquad \tau_\infty(\omega) = \sup\{t : x(\cdot, \omega) \text{ has only jumps in } (0, t)\}.$$

We call τ_∞ the *first infinity* of the M.C. Since $\{\omega : \tau_n(\omega) < t\} \in \mathfrak{F}_t$ by the definition of optionality, we have $\{\omega : \tau_\infty(\omega) < t\} \in \mathfrak{F}_t$ by (7.11). Hence τ_∞ is optional. Let

$$(7.13)\qquad L_i(t) = \mathbf{P}_i\{\tau_\infty(\omega) \leqq t\}.$$

Let $\Theta(t_1, t_2)$ denote the set $\{\omega : x(\cdot, \omega)$ has only jumps in $(t_1, t_2)\}$. For any $\Lambda \in \mathfrak{F}_t$ we have, using the optionality of τ_∞,

$$\begin{aligned}(7.14)\qquad \mathbf{P}_i\{\tau_\infty(\omega) \geqq t + t'|\Lambda; \tau_\infty(\omega) \geqq t; x(t, \omega) = j\} &\\ = \mathbf{P}_i\{\Theta(t, t + t')|x(t, \omega) = j\} &\\ = \mathbf{P}_j\{\Theta(0, t')\} = \mathbf{P}_j\{\tau_\infty(\omega) \geqq t'\}. &\end{aligned}$$

Consider a new process $\{\bar{x}_t\}$, $t \in \mathbf{T}^0$ or $\mathbf{T}$ as in $\{x_t\}$, defined as

$$(7.15)\qquad \bar{x}(t, \omega) = \begin{cases} x(t, \omega) & \text{if } t < \tau_\infty(\omega), \\ \infty & \text{if } t \geqq \tau_\infty(\omega). \end{cases}$$

We have then, if $i_\nu \in \mathbf{I}$, $1 \leqq \nu \leqq n + 1$,

$$\begin{aligned}(7.16)\qquad &\mathbf{P}\{\bar{x}(t_{n+1}, \omega) = i_{n+1}|\bar{x}(t_\nu, \omega) = i_\nu, 1 \leqq \nu \leqq n\} \\ &= \mathbf{P}\{x(t_{n+1}, \omega) = i_{n+1}; \tau_\infty(\omega) > t_{n+1}|x(t_\nu, \omega) = i_\nu, 1 \leqq \nu \leqq n; \tau_\infty(\omega) > t_n\} \\ &= \mathbf{P}\{x(t_{n+1}, \omega) = i_{n+1}; \Theta(t_n, t_{n+1})|x(t_n, \omega) = i_n\} \\ &= \mathbf{P}_{i_n}\{x(t_{n+1} - t_n, \omega) = i_{n+1}; \Theta(0, t_{n+1} - t_n)\}.\end{aligned}$$

If we put

$$(7.17)\qquad \bar{p}_{ij}(t) = \mathbf{P}_i\{x(t, \omega) = j; \Theta(0, t)\},$$

$$(7.18)\qquad \bar{p}_{i\infty}(t) = 1 - \sum_j \bar{p}_{ij}(t) = L_i(t), \qquad \bar{p}_{\infty, j}(t) = \delta_{\infty j},$$

then the last probability in (7.16) is $\bar{p}_{i_n i_{n+1}}(t_{n+1} - t_n)$ and the calculation shows that $\{\bar{x}_t\}$ is a M.C. Its state space is $\bar{\mathbf{I}}$ and its transition matrix is $(\bar{p}_{ij})$ with i and j in $\bar{\mathbf{I}}$, provided that $\mathbf{P}\{\tau_\infty(\omega) = \infty\} < 1$, or equivalently that at least one L_i is not identically zero. Otherwise $\{\bar{x}_t\}$ coincides with $\{x_t\}$.

The process $\{\bar{x}_t\}$ will be called the *minimal chain* associated with the given $\{x_t\}$. Our discussion in this section amounts to a probabilistic construction of the matrix $(\bar{p}_{ij})$, called the *minimal solution* corresponding to $Q = (q_{ij})$. We omit further properties of this matrix which will not be explicitly used below. But we note the following equation which follows from our analysis of the first discontinuity,

$$L_i(t) = \int_0^t e^{-q_i s} \sum_{j \neq i} q_{ij} L_j(t - s)\, ds. \tag{7.19}$$

Differentiating, we have

$$l_i(t) \overset{\text{def}}{=} L_i'(t) = \sum_j q_{ij} L_j(t). \tag{7.20}$$

Thus L_i has a continuous derivative. Introducing the Laplace transform

$$\hat{l}_i(\lambda) = \int_0^\infty e^{-\lambda t} l_i(t)\, dt \tag{7.21}$$

and writing $\hat{l}(\lambda)$ for the column vector $\{\hat{l}_i(\lambda)\}$, we may put (7.20) in the form

$$(\lambda I - Q)\hat{l}(\lambda) = 0. \tag{7.22}$$

8. Beyond the first infinity

We continue to assume (7.6). The first infinity τ_∞ clearly depends on the initial distribution of $\{x_t, t \in \mathbf{T}\}$. Let τ_∞^i be the restriction of τ_∞ on the set Δ_i. We rewrite (5.6) as

$$\begin{aligned} p_{jk}(t) &= \mathbf{P}_j\{\tau_\infty^j(\omega) > t;\, x(t, \omega) = k\} \\ &\qquad + \int_0^t \mathbf{P}_j\{x(t, \omega) = k \mid \tau_\infty^j(\omega) = s\}\, d\mathbf{P}_j\,\{\tau_\infty^j(\omega) \leqq s\} \\ &= \bar{p}_{jk}(t) + \int_0 \xi_{jk}(s, t) l_j(s)\, ds. \end{aligned} \tag{8.1}$$

In general $\xi_{jk}(s, t)$ is not a function of $t - s$ only, in other words [see (4) of section 4] the two fields $\mathfrak{F}_{\tau_\infty}$ and $\mathfrak{F}'_{\tau_\infty}$ are not necessarily independent. (The statement to the contrary effect on p. 235 of [1] is erroneous.) Now for an ordinary state $i \in \mathbf{I}$ we have as an easy generalization of (6.2),

$$p_{jk}(t) = {}_i p_{jk}(t) + \int_0^t p_{ik}(t - s)\, dF_{ji}\,(s). \tag{8.2}$$

If we replace the i above by ∞ and revert to our previous notation this would become, by analogy,

$$p_{jk}(t) = \bar{p}_{jk}(t) + \int_0^t \xi_k(t - s)\, dL_j\,(s). \tag{8.3}$$

Thus $\xi_{jk}(s, t)$ should not only be a function of $t - s$ only but also be independent of j. The second assertion would mean the extension of the Markov property to where $x(\tau_\infty(\omega), \omega) = \infty$, which is not asserted by the strong Markov property. The failure of (8.3) in general shows that the so-called *fictitious state* ∞ cannot be

treated like a single ordinary state, and calls for a recompactification of I. To illustrate the idea and to speak only heuristically, if only a finite number m of adjoined (fictitious) states $\infty^{(\nu)}$, $1 \leq \nu \leq m$, are needed, the situation should be as follows. To each $\infty^{(\nu)}$ corresponds an *atomic almost closed set* (see section I. 17 of [1]) $A^{(\nu)}$ of the *jump chain* $\{\chi_n, n \in \mathbf{N}\}$ in section 7, such that $x(\tau_\infty(\omega), \omega) = \infty^{(\nu)}$ iff $\chi_n(\omega) \in A^{(\nu)}$ for all sufficiently large n. Let the restriction of τ_∞ on the set $x(\tau_\infty(\omega), \omega) = \infty^{(\nu)}$ be $\tau_\infty^{(\nu)}$, and let the corresponding post-τ_∞ process be $\{y_t^{(\nu)}, t \in \mathbf{T}^0\}$. We put

$$(8.4) \qquad L_j^{(\nu)}(t) = \mathbf{P}_j\{\tau_\infty^{(\nu)}(\omega) \leq t\},$$

$$(8.5) \qquad \xi_k^{(\nu)}(t) = \mathbf{P}\{y^{(\nu)}(t, \omega) = k | \Omega_{\tau_\infty(\nu)}\}.$$

Then we should have

$$(8.6) \qquad p_{jk}(t) = \bar{p}_{jk}(t) + \sum_{\nu=1}^{m} \int_0^t \xi_k^{(\nu)}(t-s)\, dL_j^{(\nu)}(s)$$

as an improvement on (8.1). Note that each $L_j^{(\nu)}$ satisfies the same equation (7.20) as L_j and $\sum_{\nu=1}^m L_j^{(\nu)} = L_j$.

In some sense the heuristic equation (8.6) must be contained in results proved by Feller [3] by function-analytic methods. But the precise identification of the probabilistic quantities is not clear to us and in any case no probabilistic proof seems known.

If there is only one bounded nonnegative solution $\hat{l}(\lambda)$ of (7.22), apart from a scalar factor (function of λ), then $m = 1$ in (8.6) and the resulting equation (8.3) can be easily proved (see Reuter [9]). It follows from (4) of section 4 that in this case $\mathfrak{F}_{\tau_\infty}$ and $\mathfrak{F}'_{\tau_\infty}$ are independent. By (2) of section 4, we have

$$(8.7) \qquad \xi_k(s+t) = \sum_j \xi_j(s) p_{jk}(t).$$

Hence every ξ_j is continuous in $\mathbf{T}$ by lemma 1. Substituting from (8.3), we have

$$(8.8) \qquad \xi_k(s+t) = \sum_j \xi_j(s)\bar{p}_{jk}(t) + \int_0^t \xi_k(t-u)\, d_u\left[\sum_j \xi_j(s) L_j(u)\right].$$

In analogy with (5.23), we put

$$(8.9) \qquad S_\infty(\omega) = \{t : \overline{\lim_{s\to t}}\, x(s, \omega) = \infty\}.$$

We remark in this connection that $x(t, \omega)$ need not be ∞ even if $\lim_{s\uparrow t} x(s, \omega) = \infty$ or $\lim_{s\downarrow t} x(s, \omega) = \infty$ by (1.8); hence the obvious extension of (5.23) for $i = \infty$ is not adequate. Next, in analogy with (6.15) to (6.17) but for the post-τ_∞ process $\{y_t, t \in \mathbf{T}^0\}$, we put for $0 \leq s \leq t$:

$$(8.10) \qquad \delta_\infty(t, \omega) = \sup\{s : 0 \leq s \leq t, y(s, \omega) \in S_\infty(\omega)\},$$

$$(8.11) \qquad \nabla_\infty(s, t) = \mathbf{P}\{\delta_\infty(t, \omega) \leq s\} = \sum_j \xi_j(s)[1 - L_j(t-s)]$$
$$= 1 - \sum_j \xi_j(s) L_j(t-s),$$

(8.12) $$\nabla_{\infty k}(s, t) = \mathbf{P}\{\delta_\infty(t, \omega) \leqq s; y(t, \omega) = k\} = \sum_j \xi_j(s)\overline{p}_{jk}(t).$$

The quantities in (8.11) and (8.12) occur on the right side of (8.8). Letting $s \downarrow 0$, it is easy to see that the limits below exist

(8.13) $$M(t) \overset{\text{def}}{=} \lim_{s \downarrow 0} \nabla_\infty(s, s + t) = \mathbf{P}\{\delta_\infty(t, \omega) = 0\},$$

(8.14) $$\eta_k(t) \overset{\text{def}}{=} \lim_{s \downarrow 0} \nabla_{\infty k}(s, s + t) = \mathbf{P}\{\delta_\infty(t, \omega) = 0; y(t, \omega) = k\}.$$

Thus $M(t)$ is the probability that the sample function $y(\cdot, \omega)$ has only jumps in $(0, t)$, while $\eta_k(t)$ is the probability that this is so and also $y(t, \omega) = k$. Letting $s \downarrow 0$ in (8.8), we obtain

(8.15) $$\xi_k(t) = \eta_k(t) + \int_0^t \xi_k(t - u)\, dM(u).$$

This is an integral equation of the renewal type for ξ in terms of η. By definition we have

(8.16) $$\eta_k(s + t) = \sum_j \eta_j(s)\overline{p}_{jk}(t).$$

It follows by lemma 1 that $\eta_j(0)$ exist. Let

(8.17) $$\zeta_k(t) = \eta_k(t) - \sum_j \eta_j(0)\overline{p}_{jk}(t).$$

Then $\{\zeta_j\}$ satisfies the same equations (8.16) as η_j, and $\zeta_j(0) = 0$. Multiplying these equations by the "monotonicity factor" exp $(q_k t)$, see section 5, and then differentiating as in lemma 2, we obtain

(8.18) $$\zeta_k'(t) = \sum_j \zeta_j(s)\overline{p}_{jk}'(t - s)$$

for each t and almost all $s \leqq t$. Using the second system of differential equations for $(\overline{p}_{jk})$ (see section II. 17 of [1]), we conclude that

(8.19) $$\zeta_k'(t) = \sum_j \zeta_j(t)q_{jk}.$$

Passing to Laplace transforms, the last equation may be written as [compare (7.22)]

(8.20) $$\hat{\zeta}(\lambda)(\lambda I - Q) = 0.$$

The above results check with those of Reuter [9] obtained by function-analytic methods. Unfortunately it does not represent the most general case treated by Reuter, because (8.15) reduces to a trivial identity when $M = \epsilon$, or equivalently when all $\eta_k(t) \equiv 0$. However, the following positive result may be stated.

Unless the first infinity $\tau_\infty(\omega)$ is a limit point (from the right) of $S_\infty(\omega)$ with probability one, the equation (8.15) *holds nonvacuously where η and M are defined by* (8.13) *and* (8.14), *and* (8.16) *holds.*

Suppose $M(0+) - M(0-) = \beta$ where $0 \leqq \beta < 1$ and let $\tilde{M} = M - \beta$.

Solving (8.15) for ξ_k, we have

$$(1 - \beta)\, \xi_k(t) = \int_0^t \eta_k(t - u)\, dN(u), \tag{8.21}$$

where

$$N = \sum_{n=0}^{\infty} \tilde{M}^{n*} \tag{8.22}$$

in a notation similar to (5.16). If $\beta = 0$ then N is the distribution of the time between two successive points of $S_\infty(\omega)$, necessarily isolated. If $\sum_k \eta_k(0) = 1$, then this must be the case and we have the so-called "instant return from infinity" of Doob (theorem II. 19.4 of [1]). In this case we have $\zeta_k(t) = 0$ for all k. The other extreme is where $\zeta_k(t) \equiv \eta_k(t)$ for all k and only a gradual descent from infinity is possible.

The random variable $\delta_\infty(t, \cdot)$ is the last exit time from ∞ in $(0, t)$ for the post-τ_∞ process. If we consider a similar random variable $\gamma_\infty(t, \cdot)$ obtained by replacing y with x in (8.10), we are led naturally to the consideration of the following quantity, for $0 \leqq s \leqq t$,

$$\Phi_{ik}(s, t - s) \stackrel{\text{def}}{=} \mathbf{P}_i\{\gamma_\infty(\omega) \leqq s;\, x(t, \omega) = k\} = \sum_k p_{ij}(s)\bar{p}_{jk}(t - s) \tag{8.23}$$

and dually

$$\Psi_{ik}(s, t - s) \stackrel{\text{def}}{=} \mathbf{P}_i\{\tau_\infty(\omega) \geqq s;\, x(t, \omega) = k\} = \sum_j \bar{p}_{ij}(s)p_{jk}(t - s). \tag{8.24}$$

Clearly for each t, $\Phi_{ik}(s, t)$ is nondecreasing and $\Psi_{ik}(s, t)$ is nonincreasing in s. We have

$$\Phi_{ik}(0, t) = \Psi_{ik}(t, 0) = \bar{p}_{ik}(t) \tag{8.25}$$

and

$$\Phi_{ik}(t, 0) = \Psi_{ik}(0, t) = p_{ik}(t). \tag{8.26}$$

This remains true if $\bar{p}_{ij}$ in (8.23) and (8.24) is replaced by $\hat{p}_{ij}$ such that $(\hat{p}_{ij})$ is a substochastic transition matrix and

$$\hat{p}_{ij}(t) \leqq p_{ij}(t) \tag{8.27}$$

for all i and j in I. However, there are analytical difficulties if we try to differentiate $\Phi_{ik}(s, t)$ or $\Psi_{ik}(s, t)$ with respect to s. Neveu [7] overcomes these difficulties by going to Laplace transforms and we refer to his paper for further results.

REFERENCES

[1] K. L. Chung, *Markov Chains with Stationary Transition Probabilities*, Berlin, Göttingen, Heidelberg, Springer, 1960.

[2] ———, "On last exit times," *Illinois J. Math.*, Vol. 4 (1960), pp. 629–639.

[3] W. Feller, "On boundaries and lateral conditions for the Kolmogoroff differential equations," *Ann. of Math.*, Vol. 65 (1957), pp. 527–570.

[4] E. HILLE and R. S. PHILLIPS, *Functional Analysis and Semi-groups*, American Mathematical Society Colloquium Publications, Vol. 31, Providence, American Mathematical Society, 1957.

[5] W. B. JURKAT, "On the analytic structure of semi-groups of positive matrices," *Math. Z.*, Vol. 73 (1960), pp. 346–365.

[6] J. MIKUSINSKI, *Operational Calculus*, New York, Pergamon Press, 1959.

[7] J. NEVEU, "Lattice methods and submarkovian processes," *Fourth Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, University of California Press, 1961, Vol. 2, pp. 347–391.

[8] D. ORNSTEIN, "The differentiability of transition functions," *Bull. Amer. Math. Soc.*, Vol. 66 (1960), pp. 36–39.

[9] G. E. H. REUTER, "Denumerable Markov processes II," *J. London Math. Soc.*, Vol. 34 (1959), pp. 81–91.

[10] E. C. TITCHMARSH, *Introduction to the Theory of Fourier Integrals*, Oxford, Clarendon Press, 1937.

Reprinted from *Bull. Amer. Math. Sco.* **68** (1962), pp. 30–32
with permission from the American Mathematical Society.

ON THE RECURRENCE OF SUMS OF RANDOM VARIABLES

BY K. L. CHUNG[1] AND DONALD ORNSTEIN[2]

Communicated by J. L. Doob, September 19, 1961

We give a very short proof of the recurrence theorem of Chung and Fuchs [1] in one and two dimensions. This new elementary proof does not detract from the old one which uses a systematic method based on the characteristic function and yields a satisfactory general criterion. But the present method, besides its brevity, also throws light on the combinatorial structure of the problem.

Let N denote the set of positive integers, M that of positive real numbers. Let $\{X_n,\ n\in \mathrm{N}\}$ be a sequence of independent, identically distributed real-valued random vectors, and let $S_n=\sum_{\nu=1}^{n} X_\nu$. The value x is possible iff for every $\epsilon>0$ there exists an n such that $P\{|S_n-x|<\epsilon\}>0$; it is recurrent iff for every $\epsilon>0$, $P\{|S_n-x|<\epsilon$ for infinitely many $n\}=1$. It is shown in [1] that every possible value is recurrent if and only if for some $m\in\mathrm{M}$ we have

$$\sum_{n=1}^{\infty} P\{\,|S_n|\ <m\} = \infty. \tag{1}$$

---

[1] This research is supported in part by the United States Air Force Office of Scientific Research under Contract AF 49(638)-265.

[2] This research is supported in part by the NSF Grant 16434.

Note: In two dimensions $|x| = \max(|x_1|, |x_2|)$ if $x = (x_1, x_2)$. We are here concerned with obtaining sufficient conditions for the validity of (1). We state our results in two analogous propositions which correspond to one and two dimensions respectively.

PROPOSITION 1. *Let* $\{u_n(m): n \in \mathrm{N}, m \in \mathrm{M}\}$ *satisfy the following conditions:*

(i) *for each* n, $u_n(m)$ *is nonnegative, nondecreasing in* m *and* $\lim_{m\to\infty} u_n(m) = 1$;

(ii) *there exists a* $c > 0$ *such that for every positive integer* m,

$$\sum_{n=1}^{\infty} u_n(m) \leqq cm \sum_{n=1}^{\infty} u_n(1)$$

(if the left member is infinite, the inequality is taken to mean that the right member must also be infinite);

(iii) *for each* $\epsilon > 0$,

$$\lim_{n\to\infty} u_n(\epsilon n) = 1.$$

Then we have

$$\sum_{n=1}^{\infty} u_n(1) = \infty. \tag{2}$$

PROOF. Let $b \in \mathrm{N}$. We have by (i) and (ii), for integral m,

$$\sum_{n=1}^{\infty} u_n(1) \geqq \frac{1}{cm} \sum_{n=1}^{\infty} u_n(m) \geqq \frac{1}{cm} \sum_{n=1}^{bm} u_n(m) \geqq \frac{1}{cm} \sum_{n=1}^{bm} u_n\left(\frac{n}{b}\right).$$

Hence we have by (iii)

$$\sum_{n=1}^{\infty} u_n(1) \geqq \liminf_{m\to\infty} \frac{1}{cm} \sum_{n=1}^{bm} 1 = \frac{b}{c},$$

from which (1) follows since b is arbitrary, q.e.d.

REMARK. It is easy to see from the above proof that condition (iii) can be weakened, for example, to the following:

(iii*) there exists a $\delta > 0$ such that for every $\epsilon > 0$

$$\liminf_{n\to\infty} \frac{1}{n} \sum_{\nu=1}^{n} u_\nu(\epsilon\nu) \geqq \delta.$$

PROPOSITION 2. *Let* $\{u_n(m): n \in \mathrm{N}, m \in \mathrm{M}\}$ *satisfy the condition* (i) *and*

(ii$_2$) *there exists a* $c>0$ *such that for every positive integer* m

$$\sum_{n=1}^{\infty} u_n(m) \leqq cm^2 \sum_{n=1}^{\infty} u_n(1);$$

(iii$_2$) *there exist* $a>0$ *and* $d>0$ *such that*

$$u_n(m) \geqq \frac{dm}{n^{1/2}} \qquad \text{for } am^2 \leqq n.$$

Then we have (2) *again.*

PROOF. We have for $a<a'$,

$$\sum_{n=1}^{\infty} u_n(1) \geqq \frac{1}{cm^2} \sum_{am^2 \leqq n \leqq a'm^2} u_n(m) \geqq \frac{dm}{cm^2} \sum_{am^2 \leqq n \leqq a'm^2} \frac{1}{n^{1/2}}$$
$$\geqq \frac{d}{c}\left((a')^{1/2} - a^{1/2}\right)$$

for all sufficiently large m. Since a' is arbitrary, (2) follows, q.e.d.

APPLICATIONS. Take

$$u_n(m) = P\{|S_n| < m\}.$$

Condition (i) is trivially satisfied. Condition (ii) with $c=2$ or condition (ii$_2$) with $c=4$ is satisfied according as the X_n's are one-dimensional or two-dimensional. This known observation in renewal theory follows at once from the interpretation of $\sum_{n=1}^{\infty} u_n(m)$ as the expected number of entrances into the interval $(-m, m)$ by the random sequence $\{S_n, n\in N\}$. Condition (iii) is the usual normalized form of the weak law of large numbers if $E(X_n)=0$, while condition (iii$_2$) is implied by the normalized form of the central limit theorem if $E(X_n)=0$ and $E(|X_n|^2)<\infty$. Note however that here we may use the validity of these classical limit theorems as our conditions.

Let us point out that in Theorem 4a of [1] the conditions are precisely those for the validity of the weak law of large numbers in the form (iii); in Theorem 5 there, the conditions of zero mean and finite variance do imply the central limit theorem in the required form.

REFERENCE

1. K. L. Chung and W. H. J. Fuchs, *On the distribution of values of sums of random variables*, Mem. Amer. Math. Soc. No. 6 (1951), 12 pp.

STANFORD UNIVERSITY

Acta Math. **110** (1963) 19–77.
With kind permission of Springer Science and Business Media.

ON THE BOUNDARY THEORY FOR MARKOV CHAINS

BY

KAI LAI CHUNG

Stanford University, U. S. A. (1)

§ 1. Introduction and Summary

The boundary theory of Markov chains, as viewed here, is the study of essential discontinuities (viz., those which are not jumps) of the sample functions. The underlying assumptions are such that these discontinuities form a set of measure zero on the time axis and that for any given time t, the sample function will almost certainly have only jumps within an open interval containing t, reaching the boundary at both ends if at all. Thus it is a question of "how the sample curves manage to go to infinity and to come back from there" (see the preface to [1]). In Paul Lévy's terminology [9], it is a study of "fictitious states". Depending on whether the transition is to or from such a state, it is called a point on the "exit" or "entrance" boundary by Feller ([6], [7]). These ideal boundaries can be formally defined in terms of the R. S. Martin boundary theory (see [4], [5], and [8]), and the question becomes that of a suitable compactification of a discrete set, the denumerable state space of the Markov chain.

In this paper we are mainly concerned with the probabilistico-analytical aspect of the theory rather than the algebraico-topological one, if such a rough distinction may be made. Although the boundary can be defined in the general case and in more than one way, so far only the atomic part consisting of a denumerable number of boundary points has been penetrated in any sense, and substantially so only if their number is finite. It is this part which engages our attention here.

The content of this paper is most directly related to Feller's pioneering work [7]. Indeed, part of the present work arose from an effort to clarify and consolidate his results in probabilistic terms. While Feller regards his problem as one of constructing

---

(1) This research is supported in part by the Office of Scientific Research of the United States Air Force.

Markovian transition matrices out of simpler elements, here the viewpoint is that of analyzing such given matrices and their associated processes. It is perhaps a logical truism to say that a complete construction is tantamount to a complete analysis, but there is a difference in emphasis. We take the liberty to include, particularly in § 6, a number of results whose Laplace transformed versions are already in Feller's paper.

Though Feller used the language of operator theory, he has in essence created his own methods based on the resolvent equation. Reuter, in a series of papers ([12], [13] and [14]), presented the semi-group treatment of the subject and contributed to it in several respects. Neveu [11] gave a synthesis in a more general context comprising the theory of taboo states as well as boundaries. The present work has profited from the works of both authors as well as some private discussions with them and with R. S. Phillips and David Williams.

We have found it possible to derive the basic results from the first principles of probability theory together with the kind of direct methods used in [1] and [3]. Laplace transforms are employed only at a later stage. It should be mentioned that while certain analytical formulas have their "obvious" interpretations, their actual identification with probabilistic statements are not always a simple matter (see e.g. Reuter [14]). In our approach the basic quantities and their relations are obtained from considerations of the stochastic processes involved. A brief summary of the various sections will now be given.

In § 2 we give as much background material as seems feasible, though some further knowledge of the subject such as contained in §§ II. 19–20 of [1] would be necessary for a thorough understanding of the paper.

In § 3 the Martin boundary theory is reviewed. Since we can use only its atomic part its role is a rather formal one.

In § 4 the basic theorems are derived from considerations of certain martingales, and Blackwell's theorem is invoked rather than the earlier and equivalent lattice approach of Feller [6]. The crucial link is the simple but new Theorem 4.3, which as it were connects the two sides of the boundary. The rest is an application of the strong Markov property in the form given in §§ II. 8–9 of [1]. Theorem 4.6 and the open questions mentioned in its connection should serve as a test stone for any general theory of compactification of the state space of a Markov chain.

In § 5, uncomplicated probability arguments are in evidence and the fairly general Theorem 5.5 is arrived at speedily. It gives a complete description of the sample functions when there is no accumulation of boundary points in finite time and the situation may be described as being of the renewal type. Analytically, this result

already contains the first and easier case of Feller's construction. The idea of this approach was explained in [3] in the one-exit case and it is also one of the tools in Neveu [11] who found it independently.

In § 6 we use the counterpart of Feller's idea of "canonical mapping" which amounts to an integration over time in order to convert probabilities into potentials (for nonrecurrent states). Interesting, even fruitful interpretations of the results may be obtained in this light but will not be dwelt upon here. The main result is Theorem 6.3 which yields the basic relation between the transition mechanism of a given Markov chain and its "jumping" components. This must correspond to what Feller calls a lateral condition. Theorem 6.8 gives criteria for the validity of the second (forward) system of Kolmogorov differential equations which, in contrast to the first (backward) system, is not assumed throughout.

In § 7 we treat the dual chain to obtain the representation given in Theorem 7.4. This result, treated as a major consequence of our development here, is the point of departure in Feller's more algebraic theory. It must be pointed out that the dual chain is not the reversed chain (as studied in [2] in another connection) and whatever symmetry it yields is more analytic than probabilistic. However, this symmetry can be further exploited as in Neveu [11], and our lack of insistence on it may have caused some losses.

In § 8 we employ the full force of Laplace transforms as completely monotonic functions. The results may be considered as furnishing some analytical insight or hindsight on the situation. In particular, Theorem 8.3 gives a criterion for complete construction under the same conditions as in Feller [7]. From this the more explicit formulas of Feller are derived with some amendment, but a full analysis of the second case (Theorem 8.5) remains to be done.

In § 9 the one-exit case is treated in full and the results agree with those previously obtained by Reuter [13]. The connection with certain processes with independent stationary increments, discovered by Lévy [9] and analyzed by Neveu [10], is briefly mentioned.

In § 10 we give an extension of the theorem of Austin–Ornstein on the positivity of the elements of a transition matrix. While the result has only peripheral contact with the present work, it is included here for its own interest.

§ 2. Terminology and Notation

We begin with a list of symbols and conventions frequently used in this paper without further explanation. They are appreciably the same as in [1] or [3], two

major exceptions being the omission of ω wherever possible, and the use of f [1] for $\bar{p}$. Any contrary usage will be explicitly mentioned or clearly indicated by the context.

$\mathbf{N}$ is the set of nonnegative integers. The latters m, n and ν denote elements of $\mathbf{N}$.

$\mathbf{T}=[0,\infty)$; $\mathbf{T}^0=(0,\infty)$. The letters s, t, u and v denote elements of $\mathbf{T}^0$.

$\mathbf{R}$ is the set of rational numbers in $\mathbf{T}$.

$\mathbf{I}$ is a denumerable set of indices. The letters i, j and k denote elements of $\mathbf{I}$. The letters θ, θ', θ'' and $\bar{\theta}$ denote distinct objects not in $\mathbf{I}$.

In this section a statement or formula involving an unspecified element of $\mathbf{T}^0$ or $\mathbf{I}$ is meant to hold for every such element. A sequence like $\{f_i\}$ is indexed by $\mathbf{I}$; a matrix like (p_{ij}) is indexed by $\mathbf{I}\times\mathbf{I}$; a sum like $\sum_j$ is extended over $\mathbf{I}$. After this section, $\mathbf{I}$ is to be replaced by $\mathbf{I}_\theta$ (see below) in these conventions until further notice in § 6. Actually only on rare occasions does the inclusion or exclusion of θ require a careful check.

A function is real and finite valued. A function defined on $\mathbf{T}^0$ and having a right-hand limit at zero is thereby extended, together with its continuity if there is, to $\mathbf{T}$.

$$\delta_{ij}=\begin{cases}0 & \text{if } i\neq j,\\ 1 & \text{if } i=j.\end{cases}\qquad \varepsilon(t)=\begin{cases}0 & \text{if } t<0,\\ 1 & \text{if } t\geqslant 0.\end{cases}$$

$$e_q(t)=\begin{cases}0 & \text{if } t<0\\ 1-e^{-qt} & \text{if } t\geqslant 0\end{cases}\quad (0<q<\infty).$$

A (*standard*) *substochastic transition matrix is a matrix* (p_{ij}), $(i,j)\in\mathbf{I}\times\mathbf{I}$, of functions on $\mathbf{T}$ satisfying the following conditions:

$$p_{ij}(t)\geqslant 0, \tag{2.1}$$

$$\sum_j p_{ij}(s)\,p_{jk}(t)=p_{ik}(s+t), \tag{2.2}$$

$$\lim_{t\downarrow 0} p_{ij}(t)=\delta_{ij}, \tag{2.3}$$

$$\sum_j p_{ij}(t)\leqslant 1. \tag{2.4}$$

It is called *stochastic* iff equality holds in (2.4) for every i and t, and *strictly substochastic* otherwise. In the latter case we define

$$p_{i\theta}(t)\equiv 1-\sum_{j\in\mathbf{I}}p_{ij}(t),\quad p_{\theta i}(t)\equiv 0,\quad p_{\theta\theta}(t)\equiv 1 \tag{2.5}$$

[1] In honor of Feller.

and call the new matrix enlarged by the index θ the *stochastic completion* of (p_{ij}). The stochastic completion of a stochastic transition matrix is defined to be itself. Given the substochastic $\prod=(p_{ij})$, we define $\mathbf{I}_\theta$ to be $\mathbf{I}\cup\{\theta\}$ or $\mathbf{I}$ according as $\prod$ is strictly substochastic or stochastic, and define $\prod_\theta$ to be its stochastic completion.

It is known that each p_{ij} has a right-hand derivative at zero, to be denoted as follows:

$$p'_{ij}(0)=q_{ij}, \quad -q_{ii}=q_i. \tag{2.6}$$

These numbers satisfy the following relations:

$$-\infty\leqslant q_{ii}\leqslant 0, \quad 0\leqslant q_{ij}<\infty, \tag{2.7}$$

$$\sum_j q_{ij}\leqslant 0. \tag{2.8}$$

The state i is called *stable* or *instantaneous* according as $q_i<\infty$ or $q_i=\infty$; and it is *absorbing* iff $q_i=0$. The matrix (q_{ij}) will be called the *initial derivative matrix* of $\prod$ and it is said to be *conservative* iff equality holds in (2.8) for every i.

Associated with any matrix $Q=(q_{ij})$ subject to the conditions (2.7) and (2.8) are two systems of Kolmogorov differential equations:

$$z'_{ij}(t)=\sum_k q_{ik}z_{kj}(t), \tag{I_{ij}}$$

$$z'_{ij}(t)=\sum_k z_{ik}(t)\,q_{kj}. \tag{II_{ij}}$$

The *minimal solution* to both systems, first constructed by Feller, will be denoted by $\Phi=(f_{ij})$. It is a substochastic transition matrix whose initial derivative matrix is the given Q. It is minimal in this sense: if any substochastic transition matrix (p_{ij}) has the initial derivative matrix Q, then

$$f_{ij}(t)\leqslant p_{ij}(t) \tag{2.9}$$

for every i, j and t.

A (*temporally*) *homogeneous Markov chain*, or *Markov chain with stationary transition probabilities*, associated with $\mathbf{I}$ and $\prod$, is a stochastic process $\{x_t\}$, $t\in\mathbf{T}$ or $t\in\mathbf{T}^0$, on the probability triple $(\Omega,\mathfrak{F},\mathbf{P})$, having the following properties:

(i) For each t in $\mathbf{T}$ or $\mathbf{T}^0$ respectively, $x_t=x(t)$ is a discrete random variable, and the set of all possible values of all x_t is $\mathbf{I}_\theta$.

(ii) If $t_1<\ldots<t_n$, and $i_1,\ldots,i_n$ are elements of $\mathbf{I}_\theta$, then

$$\mathbf{P}\{x(t_{n+1})=i_{n+1}\,|\,x(t_\nu)=i_\nu,\ 1\leqslant\nu\leqslant n\}$$
$$=\mathbf{P}\{x(t_{n+1})=i_{n+1}\,|\,x(t_n)=i_n\}=p_{i_n i_{n+1}}(t_{n+1}-t_n).$$

A *version* of the process will be chosen to have the further properties:

(iii) For every ω in Ω,

$$x(t,\omega)=\overline{\lim_{r\downarrow t,\, r\in\mathbf{R}}}\; x(r,\omega)$$

for every t; in particular, the process is right separable with $\mathbf{R}$.

(iv) As a function of (t,ω), $x(\cdot,\cdot)$ is measurable with respect to $\mathfrak{B}\times\mathfrak{F}$ where $\mathfrak{B}$ is the usual Borel field on $\mathbf{T}$; namely, the process is Borel measurable.

From now on the process $\{x_t\}$ specified as in the above will be referred to as the given Markov chain and abbreviated as x. It is called *open* iff the parameter set is $\mathbf{T}^0$. The set $\mathbf{I}_\theta$ is called its (*minimal*) *state space*, each element of it being a *state*, and the matrix $\prod_\theta$ is called its transition matrix. The distribution of x_0, to be always concentrated on $\mathbf{I}$ rather than $\mathbf{I}_\theta$, is called its initial distribution and denoted by $\gamma=\{\gamma_i\}$, where

$$\gamma_i=\mathbf{P}\{x(0)=i\}.$$

When $\gamma_i=1$, the resulting $\mathbf{P}$ will be written as $\mathbf{P}_i$. Mathematical expectation with respect to $\mathbf{P}$ is denoted by $\mathbf{E}$, and conditional probabilities and expectations are denoted by $\mathbf{P}(\cdot\mid\cdot)$ and $\mathbf{E}(\cdot\mid\cdot)$ in the usual way.

A set like $\{\omega: x(t,\omega)=j\}$ is also written more briefly as $\{x(t)=j\}$. The *indicator function* for the set Λ is defined as follows:

$$1(\Lambda)=1_\Lambda(\omega)=\begin{cases}0 & \text{if } \omega\notin\Lambda,\\ 1 & \text{if } \omega\in\Lambda.\end{cases}$$

A property involving ω which is true for almost every ω is sometimes stated without the qualification "almost every". This can be achieved by suitably restricting the space Ω at the outset. The Borel field $\mathfrak{F}$ is assumed to be *complete* with respect to $\mathbf{P}$ and any of its subfields is supposed to be *augmented*, namely it contains all null sets. The smallest augmented Borel field with respect to which every x_s, $0\leqslant s\leqslant t$, is measurable is denoted by $\mathfrak{F}_t$.

A number of basic assumptions regarding $\prod$ or x will be gradually imposed as we proceed in the paper. They are not repeated in every theorem but any theorem given after certain assumptions have been announced is asserted to be valid under these assumptions (though they may be valid without some of them), unless exceptions are specified.

We now make the following assumption which is to hold throughout this paper.

ASSUMPTION A. *For every* $i \in \mathrm{I}$,

$$-\infty < q_{ii} < 0, \quad \sum_j q_{ij} = 0. \tag{A}$$

The first part of (A) says that every state (except θ if present) is stable and not absorbing. The second part of (A) is analytically equivalent to the validity of the first system of equations $(\hat{I}_{ij})$ for every i and j. Together they imply the following properties of the process ([1; § II.19]).

(Almost) every sample function executes an infinite sequence of jumps at the times $0 < \tau_1 < \tau_2 < \ldots$. Let $\tau_0 = 0$, and

$$\chi_n = x(\tau_n) \quad (n \in \mathrm{N}); \tag{2.10}$$

then
$$x(t) = \chi_n \quad \text{for} \quad t \in [\tau_n, \tau_{n+1}),$$

and
$$\mathbf{P}\{\tau_{n+1} - \tau_n \leqslant t \mid x(\tau_0), \ldots, x(\tau_n)\} = e_{q_{x(\tau_n)}}(t). \tag{2.11}$$

Let
$$r_{ij} = \frac{(1 - \delta_{ij})\, q_{ij}}{q_i}, \tag{2.12}$$

and $\mathrm{I}_0 = \{j : \sum_i \gamma_i r_{ij} > 0\}$. It is clear that under (A) the matrix $\mathsf{P} = (r_{ij})$ is stochastic. The stochastic process $\chi = \{\chi_n, n \in \mathrm{N})$ is a discrete parameter Markov chain with γ as its initial distribution, I_0 as its state space, and P restricted to $\mathrm{I}_0 \times \mathrm{I}_0$ as its one-step transition matrix. It is called the *jump chain* associated with x. Let

$$\tau = \lim_{n \to \infty} \tau_n; \tag{2..13}$$

then τ is a random variable which may be infinite with positive probability. It is called the *first infinity* of x.

Define further
$$\bar{x}(t, \omega) = \begin{cases} x(t, \omega) & \text{for} \quad t \in [0, \tau(\omega)), \\ \theta' & \text{for} \quad t \in [\tau(\omega), \infty). \end{cases} \tag{2.14}$$

Then the stochastic process $\bar{x} = \{\bar{x}(t), t \in \mathrm{T}\}$ is a homogeneous Markov chain with γ as its initial distribution, I_0 as its state space, and the stochastic completion of (f_{ij}) by θ' as its transition matrix. It is called the *minimal chain* associated with x. Finally, let

$$L_i(t) \equiv f_{i\theta'}(t) \equiv 1 - \sum_j f_{ij}(t), \tag{2.15}$$

then we have
$$L_i(t) = \mathbf{P}_i\{\tau \leqslant t\}. \tag{2.16}$$

Let us remark that if Q_θ and Φ_θ are the initial derivative matrix and minimal solution associated with $\prod_\theta$, they are defined on the index set $\mathbf{I}_\theta$ but not necessarily the stochastic completion of Q and Φ. Under Assumption A, it is easy to see that we have

$$\begin{aligned} &q_{i\theta}=0, \quad q_{\theta i}=0, \quad q_{\theta\theta}=1; \\ &f_{i\theta}\equiv 0, \quad f_{\theta i}\equiv 0, \quad f_{\theta\theta}\equiv 1, \quad L_\theta\equiv 0. \end{aligned} \tag{2.17}$$

It will be noticed that most formulas involving θ are either trivial or easily derived from those involving only indices in I. The extra index θ is introduced in order to employ the established formal language of probability theory which requires a total probability of one, even when we begin with a substochastic matrix.

§ 3. The Boundary

Given the matrix $\mathsf{P}=(r_{ij})$ defined in (2.12), we now choose the initial distribution γ such that $\gamma(i)>0$ for every $i\in\mathbf{I}$, so that the jump chain χ has I as its state space. Until Theorem 3.2 such terminology as "almost closed", "invariant" and "recurrent" refers to χ. According to a theorem by Blackwell (see [1; § I.17]), the set I can be decomposed as follows:

$$\mathbf{I}=\bigcup_\alpha A^\alpha, \tag{3.1}$$

where the index α ranges over a nonvoid, finite or denumerable set and where each A^α is an almost closed set, at most one of which is completely nonatomic while every other one (if any) is atomic. Furthermore if we write

$$L(A^\alpha)\doteq\limsup_n\{\chi_n\in A^\alpha\}\doteq\liminf_n\{\chi_n\in A^\alpha\}, \tag{3.2}$$

where "$\doteq$" denotes equality modulo a null set, then we have

$$\sum_\alpha \mathrm{P}\{L(A^\alpha)\}=1. \tag{3.3}$$

Without loss of generality we may suppose that the sets $L(A^\alpha)$ are disjoint. The mapping $A\to L(A)$ is a lattice isomorphism between the Borel field of equivalence classes of almost closed sets and that of equivalence classes of nonnull invariant sets. We recall that two almost closed sets are equivalent iff they differ by a transient set, and two invariant sets are equivalent iff they differ by a null set.

We define for each α:

$$\tau^\alpha=\begin{cases}\tau & \text{on } L(A^\alpha),\\ \infty & \text{on } \Omega\setminus L(A^\alpha);\end{cases} \tag{3.4}$$

and
$$\Delta = \{\tau < \infty\}, \quad \Delta^a = \{\tau^a < \infty\}. \tag{3.5}$$

It follows that $\tau^a(\omega) < \infty$ for at most one value of a on Ω, and for exactly one value of a on Δ. Note that

$$\Delta^a = \Delta \cap L(A^a), \tag{3.6}$$

and that Δ^a may be a null set; in such a case certain definitions and propositions below are vacuously true.

We now introduce the boundary for χ. In the state space $\mathbf{I}$ let the set of nonrecurrent states be $\mathbf{J}'$ and let the distinct classes of recurrent states be $\mathbf{I}_j''$ where j ranges over a possibly void, finite or denumerable set of indices $\mathbf{J}''$. Let $\mathbf{J} = \mathbf{J}' \cup \mathbf{J}''$; thus $\mathbf{J}$ is obtained from $\mathbf{I}$ by leaving the nonrecurrent states alone and identifying the states in each recurrent class as a new state. The theory of Martin boundary ([5], [8]; see also [4]) as applied to χ has the following consequences.

There exists a compact metric space $\mathbf{J}^*$ in which $\mathbf{J}$ is dense and each element of $\mathbf{J}$ is an isolated point. In other words, $\mathbf{J}^*$ is a compact metrization of $\mathbf{J}$, in which the relative topology of $\mathbf{J}$ is its natural discrete topology. The set

$$\mathbf{B} = (\mathbf{J}^* \setminus \mathbf{J}) \cup \mathbf{J}'' \tag{3.7}$$

is called the *exit boundary*, $\mathbf{J}^* \setminus \mathbf{J}$ the *nonrecurrent part* and $\mathbf{J}''$ the *recurrent part*. For almost every ω, the sequence of random variables $\{\chi_n, n \in \mathbf{N}\}$ behaves in one of the following two alternative ways:

(i) either $\chi_n(\omega)$ converges in the metric of $\mathbf{J}^*$ to a point in $\mathbf{J}''$; this happens if and only if for some j in $\mathbf{J}''$ and some m in $\mathbf{N}$, we have $\chi_n(\omega) \in I_j''$ for all $n \geqslant m$;

(ii) or $\chi_n(\omega)$ converges in the metric of $\mathbf{J}^*$ to a point in $\mathbf{J}^* \setminus \mathbf{J}$; this happens if and only if $\chi_n(\omega) \in \mathbf{J}'$ for all n in $\mathbf{N}$.

In both cases the limit, which is a random variable taking values in the boundary set $\mathbf{B}$, will be denoted by $\chi_\infty(\omega)$. Let $\mathfrak{C}$ be the topological Borel field of the metric space $\mathbf{J}^*$; the *boundary measure* μ is defined as follows:

$$\mu(C) = \mathbf{P}\{\chi_\infty \in C\} \quad (C \in \mathfrak{C}). \tag{3.8}$$

Clearly we have $\mu(\mathbf{J}') = 0$. For a singleton $\{b\} \subset \mathbf{B}$ we write $\mu(b)$ for $\mu(\{b\})$. A point b in $\mathbf{B}$ such that $\mu(b) > 0$ is called an *atomic boundary point*. Every existing recurrent class forms such a point. The set of atomic boundary points is called the (*completely*) *atomic part* of the boundary, the remaining part the (*completely*) *nonatomic part*. Either part may be void.

THEOREM 3.1. *There is a one-to-one-to-one correspondence between an atomic invariant set* Λ, *an atomic almost closed set* A, *and an atomic boundary point* b *such that*

$$\Lambda \doteq L(A) \doteq \{\chi_\infty = b\}. \tag{3.9}$$

This being so, the respective nonatomic parts are in similar correspondence.

Proof. The first correspondence in (3.9) is Blackwell's theorem cited above, and the second one is a simple consequence of a result due to Hunt [8], according to which the Borel field of all invariant sets coincides with the smallest Borel field with respect to which χ_∞ is measurable. The proof is terminated.

To proceed to the corresponding boundary for x, the time element will now be introduced. We know [1; Theorem II.19.1] that

$$\{\tau < \infty\} \doteq \left\{\sum_{n=0}^{\infty} \frac{1}{q_{\chi_n}} < \infty\right\},$$

making it manifest that the set $\{\tau < \infty\}$ is invariant and so by Hunt's result just quoted, there exists a subset $\mathbf{B}_0$ of $\mathbf{B}$, belonging to $\mathfrak{C}$, such that

$$\{\tau < \infty\} \doteq \{\chi_\infty \in \mathbf{B}_0\}. \tag{3.10}$$

Clearly $\mathbf{B}_0$ is a subset of the nonrecurrent part of the boundary; $\mathbf{B}_0$ is called the *passable part*, and $\mathbf{B}\backslash\mathbf{B}_0$ the *impassable part* of the boundary. It is important to remark that while $\mathbf{B}$ depends only on (r_{ij}), $\mathbf{B}_0$ depends on $\{q_i\}$ as well, namely it depends on (q_{ij}).

For each s in $\mathbf{T}$, let $\tau_{s,0}(\omega) = s$ and $\chi_{s,0}(\omega) = x_s(\omega)$. Let the successive times of jump of $x(\cdot, \omega)$ after s be $\{\tau_{s,n}(\omega), n \in \mathbf{N}\}$ and let

$$\chi_{s,n}(\omega) = x(\tau_{s,n}(\omega), \omega) \quad (n \in \mathbf{N}). \tag{3.11}$$

The process $\{\chi_{s,n}, n \in \mathbf{N}\}$ is called the jump chain starting at time s; it has properties similar to χ which is just the special case where $s = 0$. Let

$$\tau_{s,\infty} = \lim_{n\to\infty} \tau_{s,n}, \quad \chi_{s,\infty} = \lim_{n\to\infty} \chi_{s,n}, \tag{3.12}$$

the latter limit being again in the topology of $\mathbf{J}^*$. We shall say that after the given time s, the boundary is first reached at time $\tau_{s,\infty}$ and at the point $\chi_{s,\infty}$. Given a subinterval S of $\mathbf{T}$, the boundary is reached in S at b iff there is an s in S such that $\tau_{s,\infty} \in S$ and $\chi_{s,\infty} = b$. Note that the state space of $\{\chi_{s,n}\}$, as well as the corre-

sponding boundary measure $\mu_s(C)=\mathbf{P}\{\chi_{s,\infty}\in C\}$, may vary with s. Since for each given s, almost every sample function is constant in an open interval containing s, it is sufficient to consider all jump chains starting at rational times. More precisely, for almost all ω, all $\tau_{r,\infty}(\omega)$ and $\chi_{r,\infty}(\omega)$ are well defined simultaneously for all r in $\mathbf{R}$; and for each fixed s, we have for almost every ω (the exceptional null set depending on s):

$$\tau_{s,\infty}(\omega)=\lim_{r\to s}\tau_{r,\infty}(\omega),\quad \chi_{s,\infty}(\omega)=\lim_{r\to s}\chi_{r,\infty}(\omega). \tag{3.13}$$

However, it is false that for almost every ω, there is a first time that the boundary is reached after every (generic, not fixed) t. Indeed this may be false for t equal to the first infinity $\tau(\omega)$, and here lies much of the difficulty of the theory.

The boundary concepts given in this section are maximal ones relative to a given matrix P or Q. A smaller boundary can be defined relative, in addition, to a given initial distribution γ. By choosing an everywhere positive γ to begin with we have in effect covered all possible choices of γ, and so in particular included the boundary of $\{\chi_{s,n}\}$ for every s.

We conclude this section by a description of the set of states from which the nonrecurrent and passable part of boundary can not be reached. A sample path beginning at such a state will either reach the recurrent part of the boundary in finite time or remain indefinitely in some almost closed set, approaching the impassable part of the boundary as times goes to infinity. Let

$$Z=\{i\in\mathbf{I}:L_i\equiv 0\}. \tag{3.14}$$

In the following, the notions "stochastically closed" and "recurrence" will be prefixed by $\prod$ or Φ according to the transition matrix they refer to.

THEOREM 3.2. *The set Z is the set of i such that*

$$\mathbf{P}_i\{\tau=\infty\}=1. \tag{3.15}$$

It is $\prod$-stochastically closed and contains all Φ-recurrent states.

Proof. The first assertion follows at once from (2.16). Next, if $i\in Z$, then by (2.9),

$$1=\sum_j f_{ij}(t)\leqslant\sum_j p_{ij}(t)\leqslant 1; \tag{3.16}$$

hence $f_{ij}\equiv p_{ij}$. It follows from the definition (2.15) that

$$L_i(s+t)-L_i(s)=\sum_j f_{ij}(s)\,L_j(t); \tag{3.17}$$

hence if $i \in Z$, and $f_{ij}(s) = p_{ij}(s) > 0$ for some s, then $L_j \equiv 0$, proving that Z is $\prod$-stochastically closed. Furthermore we deduce from (3.17) that for every m and s:

$$1 \geqslant \sum_{n=0}^{m} [L_i((n+1)s) - L_i(ns)] \geqslant \sum_{n=0}^{m} f_{ii}(ns) L_i(s). \tag{3.18}$$

If i is Φ-recurrent, then
$$\sum_{n=0}^{\infty} f_{ii}(ns) = \infty. \tag{3.19}$$

It follows from (3.18) and (3.19) that $L_i(s) = 0$. This being true for any s, we conclude that $i \in Z$.

An alternative proof of the last part of the theorem is as follows. By a fundamental result on stable states ([1; Theorem 5.7]), the number of disjoint i-intervals for $x(t, \omega)$ is finite in any finite subinterval of T for almost every ω. Consequently, the total number of disjoint i-intervals for the minimal chain is finite whenever $\tau(\omega) < \infty$. On the other hand, if i is Φ-recurrent (and not absorbing), this number must be infinite and so (3.15) must hold, hence $i \in Z$.

Remark. It is possible that i is $\prod$-recurrent and yet $i \notin Z$. We need only take an infinite number of independent copies of an ascending escalator, hitched onto one another (see [1; § II.20]). Every state is Φ-nonrecurrent and $\prod$-recurrent in the resulting chain, and Z is void.

COROLLARY TO THEOREM 3.2. $Z = \mathrm{I}$ *if and only if the passable part of the boundary is void, or* $\prod = \Phi$.

§ 4. Fundamental Theorems

Recalling (3.4) we put
$$L_i^a(t) = \mathbf{P}_i\{\tau^a \leqslant t\}; \tag{4.1}$$

then $L_i^a(t)$ is the probability, starting from i, that the first infinity is reached no later than at time t, and that the jump chain finishes by remaining in the almost closed set A^a. We have clearly

$$L_i = \sum_a L_i^a; \tag{4.2}$$

and the analogue of (3.17) holds:

$$L_i^a(s+t) - L_i^a(s) = \sum_j f_{ij}(s) L_j^a(t), \text{(1)} \tag{4.3}$$

(1) By our convention $i \neq \theta$ and $j \neq \theta$, but we can in virtue of (2.17) extend this and similar formulas to cover the index θ.

this time from its probabilistic meaning. By a general analytical lemma ([3; Lemma 2]), the equation (4.3) implies that each L_i^a has a continuous derivative l_i^a in T satisfying

$$l_i^a(s+t) = \sum_j f_{ij}(s)\, l_j^a(t). \tag{4.4}$$

Furthermore, the Kolmogorov equation ($\mathrm{I}_{ii'}$) for $f_{ii'}$ (see (2.15)) reduces to

$$l_i^a(t) = \sum_j q_{ij} L_j^a(t). \tag{4.5}$$

Since $L_j^a(0)=0$ it follows that

$$l_i^a(0) = 0. \tag{4.6}$$

THEOREM 4.1. *We have for every t: $0 < t \leqslant \infty$, with probability one:*

$$\lim_{n\to\infty} L_{\chi_n}^a(t) = 1(\Delta^a). \tag{4.7}$$

Proof. We have ([1; § II.19])

$$\mathbf{E}\{\tau - \tau_n \,|\, \chi_n\} = \sum_{m=n}^{\infty} q_{\chi_m}^{-1}, \tag{4.8}$$

where the series converges on Δ and diverges on $\Omega\setminus\Delta$. By Lévy's zero-or-one law, we have on Δ^a:

$$\lim_{n\to\infty} \mathbf{P}\{\tau = \tau^a \,|\, \chi_n\} = 1,$$

and consequently

$$\lim_{n\to\infty} \mathbf{E}\{\tau^a - \tau_n \,|\, \chi_n\} = 0.$$

Since

$$1 - L_{\chi_n}^a(t) = \mathbf{P}\{\tau^a - \tau_n > t \,|\, \chi_n\} \leqslant t^{-1}\, \mathbf{E}\{\tau^a - \tau_n \,|\, \chi_n\}, \tag{4.9}$$

it follows that the first member in (4.9) converges to zero as $n\to\infty$. On the other hand, by the same law we have on $\Omega\setminus\Delta^a$:

$$\lim_{n\to\infty} \mathbf{P}\{\tau^a = \infty \,|\, \chi_n\} = 1.$$

Hence it follows from the first relation in (4.9) that its first member converges to one as $n\to\infty$. Thus (4.7) is proved for $0<t<\infty$. This trivially implies (4.7) for $t=\infty$ on Δ^a, which in turn implies the same on $\Omega\setminus\Delta^a$ by (4.2). Theorem 4.1 is proved.

COROLLARY. *For each a such that $\mathbf{P}\{\Delta^a\}>0$, there exists a sequence of states $\{i_n\}$ such that for every b:*

$$\lim_{n\to\infty} L_{i_n}^b(\cdot) = \delta^{ab}\,\varepsilon(\cdot).$$

In particular if
$$\hat{l}_i^b(\lambda)=\int_0^\infty e^{-\lambda t}\,dL_i^b(t),$$
then
$$\lim_{n\to\infty}\hat{l}_{i_n}^b(\lambda)=\delta^{ab}.$$

Let $C\in\mathfrak{C}$, then $\{\chi_\infty\in C\}$ is an invariant set for χ, hence if we set $U_i=\mathbf{P}_i\{\chi_\infty\in C\}$, we have
$$U_i=\sum_j r_{ij}U_j \quad\text{or}\quad \sum_j q_{ij}U_j=0. \tag{4.10}$$

Define the function $U_i(\cdot)$ on $\mathbf{T}$ as follows:
$$U_i(t)=U_i-\sum_j f_{ij}(t)\,U_j, \tag{4.11}$$
then we have
$$U_i(t)=\mathbf{P}_i\{\chi_\infty\in C;\ \tau\leqslant t\}. \tag{4.12}$$

It follows from (4.11) or (4.12) that
$$U_i(s+t)-U_i(s)=\sum_j f_{ij}(s)\,U_j(t); \tag{4.13}$$
hence by the lemma cited after (4.3), each $U_i(\cdot)$ has a continuous derivative $u_i(\cdot)$ satisfying
$$u_i(s+t)=\sum_j f_{ij}(s)\,u_j(t)\quad (s>0,\ t>0). \tag{4.14}$$

If C is a subset of $\mathbf{B}\setminus\mathbf{B}_0$, then $U_i(\cdot)\equiv 0$ for every i by (4.12), and conversely. Otherwise, $u_i(t)>0$ for some i and $t>0$ (see the Appendix for a stronger result). If C is a passable atomic boundary point corresponding to A^a, then $U_i(\cdot)$ reduces to $L_i^a(\cdot)$.

A set of nonnegative functions $\{u_i(\cdot)\}$ with $u_i(0)=0$ for every i and satisfying (4.14) will be called an *exit solution* for Φ. If the u_i's are nonnegative and satisfy (4.14), and we set
$$\bar{u}_i(t)=u_i(t)-\sum_j f_{ij}(t)\,u_j(0),$$
then $\{\bar{u}_i(\cdot)\}$ is an exit solution for Φ.

We now make our second basic assumption:

Assumption B. *The passable part of the boundary is nonvoid and completely atomic.*

We shall denote these atomic boundary points by $\{\infty^a,\ a\in\mathbf{A}\}$, where $\mathbf{A}$ is a nonvoid, finite or denumerable index set. We have thus
$$\{\tau<\infty\}=\bigcup_{a\in\mathbf{A}}\{\tau^a<\infty\}=\bigcup_{a\in\mathbf{A}}\{\tau<\infty;\ \chi_\infty=\infty^a\}.$$

Theorem 4.2. *Under Assumption B, every exit solution satisfying*
$$\int_0^\infty u_i(t)\,dt\leqslant 1 \tag{4.15}$$

is given by

$$u_i(t) = \sum_{a \in \mathbf{A}} c^a l_i^a(t), \tag{4.16}$$

where $0 \leqslant c^a \leqslant 1$ *for every* a *in* $\mathbf{A}$. *Furthermore the representation* (4.16) *is unique.*

Proof. Let $\{u_i(\cdot)\}$ satisfy (4.14) and (4.15). Since the Kolmogorov equation (I_{ij}) for Φ is equivalent to

$$f_{ij}(s) = \sum_{k \neq i} \int_0^s e^{-q_i(s-v)} q_{ik} f_{kj}(v)\, dv + \delta_{ij}, \tag{4.17}$$

we have, upon substitution into (4.14):

$$u_i(s+t) - u_i(t) = \sum_{k \neq i} \int_0^s e^{-q_i(s-v)} q_{ik} u_k(v+t)\, dv.$$

It is easy to see that we can let $t \downarrow 0$ termwise under the integral; and integrating the resulting equation over $(0, \infty)$, we obtain

$$\int_0^\infty u_i(s)\, ds = \sum_{k \neq i} q_i^{-1} q_{ik} \int_0^\infty u_k(v)\, dv.$$

Hence if we set $U_i = \int_0^\infty u_i(s)\, ds$, $\{U_i\}$ is a solution of (4.10) with $0 \leqslant U_i \leqslant 1$. By a theorem due to Blackwell ([1; § I.17]), to such a solution there corresponds an invariant function φ with $0 \leqslant \varphi \leqslant 1$ such that $U_i = \mathrm{E}_i(\varphi)$. Now decompose φ as follows:

$$\varphi = \sum c^a 1(\Lambda^a) + \varphi^0 \quad (0 \leqslant c^a \leqslant 1),$$

where $\Lambda^a = L(A^a)$ in (3.2), and the sum is over the disjoint atomic invariant sets, φ^0 being the remainder which vanishes on the atomic invariant sets. We have, correspondingly,

$$U_i = \sum c^a \mathrm{P}_i(\Lambda^a) + \mathrm{E}_i(\varphi^0),$$

and using (4.11), the discussion after (4.14), and Assumption B:

$$U_i(t) = \sum_{a \in \mathbf{A}} c^a L_i^a(t) = \int_0^t \sum_{a \in \mathbf{A}} c^a l_i^a(s)\, ds.$$

Upon differentiation we obtain (4.16) a. e. Since $\sum_a l_i^a(t) = l_i(t)$, the series converges uniformly in every compact interval of $\mathbf{T}$ by Dini's Theorem. It follows that both members of (4.16) are continuous and so the equation holds for every t.

Suppose $u_i(\cdot)$ has another representation of the form (4.16) with c^a replaced by d^a; it follows that

$$\sum_{a\in \mathbf{A}} c^a L_t^a(t) = \sum_{a\in \mathbf{A}} d^a L_t^a(t).$$

Applying the Corollary to Theorem 4.1, we have $c^a = d^a$. Thus the representation is unique and Theorem 4.2 is proved.

By the same argument, we see that the set of exit solutions $\{l_i^a(\cdot)\}$, $a\in \mathbf{A}$, is a linearly independent set. The conclusion of Theorem 4.2 may be expressed by saying that this set is the *extreme base* of the space of exit solutions.

From now on an unspecified super-index a or b denotes an element of $\mathbf{A}$ and an unspecified sum over it extends over $\mathbf{A}$.

For terminology relevant to optionality see [1; § II. 8–9]. In particular, if τ is optional. $\mathfrak{F}_\tau$ and $\mathfrak{F}'_\tau$ denote respectively the pre-τ and post-τ fields.

LEMMA. *Each τ^a is an optional random variable.*

Proof. For each $n\in \mathbf{N}$, τ_n is optional, as can be seen by induction on n. Next, let $\tau_{n,m} = [m\tau_n + 1]/m$ for every positive integer m. Then $\tau_{n,m}$ is rationally valued and optional, and $\tau_{n,m} < t$ for all sufficiently large m on the set $\{\tau < t\}$. These facts imply that $x(\tau_{n,m})$ is measurable with respect to the pre-τ field $\mathfrak{F}_\tau$ for large m. Since almost every sample function is constant in a right-hand neighborhood of every τ_n, by the basic property of a stable state, we have

$$\lim_{m\to\infty} x(\tau_{n,m}) = x(\tau_n)$$

with probability one by the specification (iii) in § 2 of x. Hence every $x(\tau_n)$ is measurable with respect to $\mathfrak{F}_\tau$. Now

$$\{\tau^a < t\} = \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \{\tau_n < t;\, x(\tau_n)\in A^a\} \in \mathfrak{F}_t.$$

The lemma is proved.

Under Assumption B, if γ is positive everywhere as in § 3, then $\mathbf{P}(\Delta^a) > 0$ for every a in $\mathbf{A}$. For an arbitrary γ, some $\mathbf{P}(\Delta^a)$ may be zero. In what follows, we shall tacitly suppose that $\mathbf{P}(\Delta^a) > 0$ in a discussion involving Δ^a.

The *post-τ^a process* $x^a = \{x_t^a, t\in \mathbf{T}^0\}$ on Δ^a is defined as follows:

$$x_t^a = x^a(t) = x(\tau^a + t). \tag{4.18}$$

By the *strong Markov property* as discussed in [1; § II. 9], x^a is an open, homogeneous Markov chain on the probability triple $(\Delta^a, \mathfrak{F}^a, \mathbf{P}^a)$, where $\mathfrak{F}^a = \mathfrak{F}\cap \Delta^a$, $\mathbf{P}^a(\cdot) = \mathbf{P}(\cdot \mid \Delta^a)$, with a certain subset $\mathbf{I}^a$ of $\mathbf{I}_\theta$ as its state space and with the restriction of $\prod_\theta$ to $\mathbf{I}^a\times \mathbf{I}^a$ as its transition matrix. Properties corresponding to (iii) and (iv) in § 2 also hold.

The process x^a can be extended to the parameter set $\mathbf{T}$ on Δ^a if and only if $x^a(0) = x(\tau^a) \in \mathbf{I}_\theta$ almost everywhere on Δ^a.

The next few theorems and their corollaries form the probabilistic basis of the present investigation.

THEOREM 4.3. *For each a in $\mathbf{A}$, each j in $\mathbf{I}_\theta$ and each $t > 0$, there exists a number $\xi_j^a(t)$ such that*

$$\lim_{n\to\infty} \mathrm{P}\{\Delta^a; x^a(t) = j \mid \chi_n\} = 1(\Delta^a)\xi_j^a(t). \tag{4.19}$$

Proof. For each n and t, we have

$$\tau^a + t = \lim_{m\to\infty}\left\{\tau_n + \frac{[m(\tau^a + t - \tau_n) + 1]}{m}\right\}.$$

It follows from an argument similar to that in the lemma above that $x(\tau^a + t)$ is measurable with respect to the post-τ_n field $\mathfrak{F}'_{\tau_n}$ and hence with respect to $\bigcap_n \mathfrak{F}'_{\tau_n}$. This fact and the fact that χ is Markovian imply that the limit in (4.19) exists by the martingale convergence theorem, and being the limit it is *ipso facto* an invariant function for χ. On $\Omega\backslash\Delta^a$ it is zero by Lévy's zero-or-one law. In general it is constant on each atomic invariant set $\Lambda^a = L(A^a)$. This constant is the $\xi_j^a(t)$ in (4.19). Theorem 4.3 is proved.

COROLLARY. *We have* $\mathbf{P}\{x^a(t) = j \mid \Delta^a\} = \xi_j^a(t)$,

so that $\{\xi_j^a(t)\}$ is the absolute distribution of the post-τ^a process at time t.

It follows that
$$\sum_{i\in I^a} \xi_i^a(t) = 1, \tag{4.20}$$

$$\sum_{i\in \mathbf{I}^a} \xi_i^a(s)p_{ij}(t) = \xi_j^a(s+t). \tag{4.21}$$

In particular by a general analytical lemma ([3; Lemma 1]), each ξ_i^a is continuous in $\mathbf{T}$.

Define $\mathfrak{F}_\tau^- = \mathsf{V}_n\mathfrak{F}_{\tau_n}$. Note that in general $\mathfrak{F}_\tau^-$ is a proper subfield of $\mathfrak{F}_\tau$, but τ is measurable with respect to $\mathfrak{F}_\tau^-$.

THEOREM 4.4. *For each a in $\mathbf{A}$, if $\mathrm{M} \in \mathfrak{F}_\tau^-$ and $\mathrm{M}' \in \mathfrak{F}'_\tau$, then*

$$\mathbf{P}\{\mathrm{MM}' \mid \Delta^a\} = \mathbf{P}\{\mathrm{M} \mid \Delta^a\}\mathbf{P}\{\mathrm{M}' \mid \Delta^a\}. \tag{4.22}$$

Proof. Let $n \in \mathrm{N}$ and $\mathrm{M} \in \mathfrak{F}_{\tau_n}$. If $n < m$, then $\mathfrak{F}_{\tau_n} \subset \mathfrak{F}_{\tau_m}$ trivially and $\mathfrak{F}'_\tau \subset \mathfrak{F}'_{\tau_m}$ by

the argument at the beginning of the preceding proof. Hence $M \in \mathfrak{F}_{\tau_m}$ and $\{x^a(t)=j\} \in \mathfrak{F}'_{\tau_m}$. Applying the strong Markov property to τ_m, we have [1; Theorem II. 9.3]:

$$\mathbf{P}\{M;\ \Delta^a;\ x^a(t)=j \mid \chi_m\} = \mathbf{P}\{M \mid \chi_m\}\,\mathbf{P}\{\Delta^a;\ x^a(t)=j \mid \chi_m\}.$$

Integrating over Ω and letting $m\to\infty$, we have

$$\mathbf{P}\{M;\ \Delta^a;\ x^a(t)=j\} = \int_{\Omega} \lim_{m\to\infty} \mathbf{P}\{M \mid \chi_m\}\,\mathbf{P}\{\Delta^a;\ x^a(t)=j \mid \chi_m\}\,d\mathbf{P}. \tag{4.23}$$

Now if $\mathfrak{I}$ denotes the invariant field for χ, then $\Delta^a \in \mathfrak{I}$ and by the Markovian property of χ and a simple martingale convergence theorem:

$$\lim_{m\to\infty} \mathbf{P}\{M \mid \chi_m\} = \mathbf{P}\{M \mid \chi_m, \chi_{m+1}, \ldots\} = \mathbf{P}\{M \mid \mathfrak{I}\};$$

consequently

$$\int_{\Delta^a} \lim_{m\to\infty} \mathbf{P}\{M \mid \chi_m\}\,d\mathbf{P} = \mathbf{P}\{M;\ \Delta^a\}. \tag{4.24}$$

Using (4.19) and (4.24) in (4.23), we obtain

$$\mathbf{P}\{M;\ x^a(t)=j \mid \Delta^a\} = \mathbf{P}\{M \mid \Delta^a\}\,\xi_j^a(t).$$

Applying the strong Markov property to τ^a+t, we obtain furthermore that if $0<t=t_1<\cdots<t_l$, then

$$\begin{aligned}
&\mathbf{P}\{M;\ x^a(t_\nu)=j_\nu,\ 1\leqslant \nu\leqslant l \mid \Delta^a\}\\
&= \mathbf{P}\{M \mid \Delta^a\}\,\xi_{j_1}^a(t_1)\prod_{\nu=1}^{l-1} p_{j_\nu j_{\nu+1}}(t_{\nu+1}-t_\nu)\\
&= \mathbf{P}\{M \mid \Delta^a\}\,\mathbf{P}\{x^a(t_\nu)=j_\nu,\ 1\leqslant \nu\leqslant l \mid \Delta^a\}.
\end{aligned}$$

This being true for arbitrary t_ν's and j_ν's we conclude that (4.22) holds for every $M \in \mathfrak{F}_{\tau_n}$ and $M' \in \mathfrak{F}'_\tau$; since n is arbitrary it holds also for every $M \in \bigvee_n \mathfrak{F}_{\tau_n} = \mathfrak{F}_\tau^-$. Theorem 4.4 is proved.

THEOREM 4.5. *For almost every ω in Δ^a, $t_n(\omega)\to t$ implies*

$$\lim_{n\to\infty} p_{\chi_{r,n}(\omega),j}(t_n(\omega)) = \xi_j^a(t) \tag{4.25}$$

for every r in $\mathbf{R}$ *and j in* $\mathbf{I}_0$.

Proof. We first prove (4.25) with $r=0$ and all $t_n(\omega)$ equal to a fixed t, and with the exceptional null set possibly varying with t. Given $\varepsilon>0$, there exist Γ and m such

that $\Gamma \subset \Delta^a$ and $\mathbf{P}(\Delta^a \backslash \Gamma) < \varepsilon$, and such that if $\omega \in \Gamma$ and $n > m$, we have

$$\tau^a(\omega) - \tau_n(\omega) < \varepsilon. \tag{4.26}$$

The basic property of the stable state j for the post-τ_n process implies that on Γ:

$$\mathbf{P}\{\Delta^a; x^a(t) = j \mid \chi_n\} \geq \mathbf{P}\{x(\tau_n + t) = j \mid \chi_n\} e^{-q_j \varepsilon} = p_{\chi_n, j}(t) e^{-q_j \varepsilon}.$$

Letting $n \to \infty$, then $\varepsilon \to 0$, we obtain from Theorem 4.3:

$$\xi_j^a(t) \geq \overline{\lim_{n \to \infty}}\, p_{\chi_n, j}(t)$$

almost everywhere on Δ^a. Similarly we have if $\varepsilon < t$,

$$\mathbf{P}\{\Delta^a; x^a(t-\varepsilon) = j \mid \chi_n\} e^{-q_j \varepsilon} \leq \mathbf{P}\{x(\tau_n + t) = j \mid \chi_n\} = p_{\chi_n, j}(t).$$

Passing to limits as before, we obtain

$$\xi_j^a(t) \leq \underline{\lim_{n \to \infty}}\, p_{\chi_n, j}(t)$$

almost everywhere on Δ^a. Thus (4.25) is true in this case. Applying this a doubly denumerable number of times, we infer that for almost every ω in Δ^a, and for a fixed sequence $\{r_m\}$ converging to zero,

$$\lim_{n \to \infty} p_{\chi_{r,n}(\omega), j}(r_m) = \xi_j^a(r_m) \tag{4.27}$$

for every r, j and m. For any ω for which (4.27) holds, and for which furthermore $\chi_{r,n}(\omega) \in \mathbf{I}_\theta$ for every r and n, we now show that the stronger (4.25) also holds, as follows. Let $\chi_{r,n}(\omega) = i_n$, $t_n(\omega) = t_n$. For any $t > 0$, there exists m such that $t > r_m$; hence $t_n > r_m$ for all sufficiently large n. We have by Fatou's lemma and (4.21):

$$\begin{aligned} \underline{\lim_{n \to \infty}}\, p_{i_n, j}(t_n) &\geq \sum_i \underline{\lim_{n \to \infty}}\, p_{i_n, i}(r_m) p_{ij}(t_n - r_m) \\ &= \sum_i \xi_i^a(r_m) p_{ij}(t - r_m) = \xi_j^a(t). \end{aligned} \tag{4.28}$$

Consequently by (4.20)

$$1 \geq \sum_{j \in \mathbf{I}_\theta} \underline{\lim_{n \to \infty}}\, p_{i_n, j}(t_n) \geq \sum_{j \in \mathbf{I}_\theta} \xi_j^a(t) = 1.$$

Hence equality holds in (4.28), and this remains true if $\{i_n\}$ is replaced by a subsequence. Therefore the $\underline{\lim}$ in (4.28) may be replaced by lim and Theorem 4.5 is proved.

As a consequence of Theorem 4.5, we shall prove a property of sample functions regarding the time set on which the boundary is reached. The corresponding property [1; Theorem II. 6.1], first proved by Doob for an ordinary (vs. a fictitious) state, is a major result in the theory of Markov chains. Two passable atomic boundary points ∞^a and ∞^b, $a \neq b$, are said to be *indistinguishable* iff the corresponding post-τ^a and post-τ^b processes have the same finite-dimensional distributions; otherwise they are *distinguishable*. If ∞^a and ∞^b are indistinguishable, they can be "merged" as follows. Define

$$\{\infty^{a\cup b}\} = \{\infty^a\} \cup \{\infty^b\}; \qquad \Delta^{a\cup b} = \Delta^a \cup \Delta^b;$$

$$\tau^{a\cup b} = \begin{cases} \tau^a & \text{on} \quad \Delta^a, \\ \tau^b, & \text{on} \quad \Delta^b, \end{cases} \qquad x^{a\cup b}(t) = \begin{cases} x^a(t) & \text{on} \quad \Delta^a, \\ x^b(t), & \text{on} \quad \Delta^b; \end{cases}$$

$$L_i^{a\cup b}(t) \equiv L_i^a(t) + L_i^b(t), \qquad \xi^{a\cup b}(t) \equiv \xi^a(t) \equiv \xi^b(t);$$

and treat the union of the two atomic boundary points as if they were one.

For each ω, let the set of t for which there is an $s < t$ such that $\tau_{s,\infty}(\omega) = t$ and $\chi_{s,\infty}(\omega) = \infty^a$ be denoted by $S_{\infty^a}(\omega)$. This is the time set on which the sample function $x(\cdot, \omega)$ reaches the boundary at the point ∞^a. The union of $S_{\infty^a}(\omega)$ over a in A may be denoted by $S_\infty^-(\omega)$ and is the time set on which $x(\cdot, \omega)$ reaches the boundary (under Assumptions A and B). Note that $t \in S_\infty^-(\omega)$ does not imply $x(t, \omega) = \infty$, according to the specification (iii) in § 2.

THEOREM 4.6. *If ∞^a and ∞^b are distinguishable, then for almost every ω, no t is a left-hand [or right-hand]*(1) *limit point of both $S_{\infty^a}(\omega)$ and $S_{\infty^b}(\omega)$.*

Proof. Let Ω_0 with $\mathbf{P}(\Omega_0) = 1$ be so chosen that (i) every stable state is taken in an open set; (ii) for every j, s and t where $s < t$ the martingale

$$p_{x(r),j}(t-r)$$

as $r \downarrow s$ or $r \uparrow s, r \in \mathbf{R}$, has a unique limit; (iii) Theorem 4.5 holds for every a. The second stipulation is possible by the regularity properties of the sample functions of a martingale (see [1; p. 153]). We now show that if for some ω_0 in Ω_0, and an s in $\mathbf{T}^0$, both $S_{\infty^a}(\omega_0)$ and $S_{\infty^b}(\omega_0)$ intersect $(s, s+\delta)$ for arbitrarily small δ, then $\xi^a(\cdot) \equiv \xi^b(\cdot)$. A similar proof holds for $(s-\delta, s)$.

By hypothesis, for every δ there exists an r in $\mathbf{R}$ such that

(1) t is a left-hand or right-hand limit point of S according as $(t, t+\delta) \cap S \neq 0$ or $(t-\delta, t) \cap S \neq 0$ for every $\delta > 0$.

$$\tau_{r,\infty} \in (s, s+\delta) \quad \text{and} \quad \chi_{r,\infty} = \infty^a.$$

Hence by (i) and (iii) above, there exists a sequence $s_n \in \mathbf{R}$, $s_n \uparrow \tau_{r,\infty}$ such that if $t > s + \delta$,

$$\lim_{n\to\infty} p_{x(s_n),j}(t - s_n) = \xi_j^a(t - \tau_{r,\infty}).$$

Since δ is arbitrary and ξ_j^a is continuous, this implies that there exists a sequence $r_n \in \mathbf{R}$, $r_n \downarrow$ $s$such that if $t > s$,

$$\lim_{n\to\infty} p_{x(r_n),j}(t - r_n) = \xi_j^a(t - s).$$

A similar relation holds for another sequence $r_n' \in \mathbf{R}$, $r_n' \downarrow s$, and ξ_j^b instead of ξ_j^a. Therefore by (ii) above,

$$\xi_j^a(t - s) = \xi_j^b(t - s).$$

This being true for every $t > s$, we have $\xi_j^a(\cdot) \equiv \xi_j^b(\cdot)$, proving the theorem.

Remark. A point of jump t is a right-hand limit point of some $S_i(\omega)$ and a left-hand limit point of a distinct $S_j(\omega)$. Without assuming that the states are stable, i and j may be instantaneous if the unilateral limits are replaced by lower limits. From this we surmise that the unilaterality stipulation in the theorem is necessary, though we are not giving a specific counterexample, nor one to show that the distinguishability is also necessary.

§ 5. The First Approach

The starting point of the analysis of probabilities of transition *to* and from the boundary is the following result.

THEOREM 5.1. *Under Assumptions A and B, we have*

$$p_{ij}(t) = f_{ij}(t) + \sum_a \int_0^t l_i^a(s) \xi_j^a(t-s)\, ds. \tag{5.1}$$

Proof. We have

$$\mathbf{P}_i\{x_t = j\} = \mathbf{P}_i\{\tau > t;\, x_t = j\} + \sum_a \mathbf{P}_i\{\tau^a \leqslant t;\, x_t = j\}. \tag{5.2}$$

The first term on the right side is, according to (2.14), $\mathbf{P}_i\{\bar{x}_t = j\} = f_{ij}(t)$. Next, the conditional independence asserted in Theorem 4.4 implies, by [1; Theorem II. 9.4], that

$$\mathbf{P}_i\{\tau^a \leqslant t;\, x_t = j\} = \int_0^t \xi_j^a(t-s)\, dL_i^a(s) = \int_0^t l_i^a(s)\, \xi_j^a(t-s)\, ds.$$

Substituting into (5.2) we obtain (5.1). Theorem 5.1 is proved.

For any subinterval (s, t) of $\mathbf{T}^0$, we define

$$0^a(s,t)=\{\omega: x^a(\cdot,\omega) \text{ does not reach the boundary in } (s,t)\}$$
$$=\{\omega: x^a(\cdot,\omega) \text{ has only jumps in } (s,t)\}. \tag{5.3}$$

It is immaterial whether the interval (s, t) is open or closed, provided that a null set can be ignored. Now we define for each t:

$$\delta_t^a(\omega)=\delta^a(t,\omega)=\inf\{s: 0\leqslant s\leqslant t;\ \omega\in 0^a(s,t)\}. \tag{5.4}$$

It is easy to see that δ_t^a is a random variable with a continuous distribution, which will be given shortly. We call $\delta^a(t,\omega)$ the *last exit from the boundary before time t for the sample function* $x^a(\cdot,\omega)$. By definition $\delta^a(t,\omega)$ either belongs to or is a right-hand limit point of the set

$$S_\infty^-(\omega)=\bigcup_{a\in\mathbf{A}} S_{\infty a}(\omega),$$

but it may be the left-hand endpoint of a stable interval. Even if $\mathbf{A}$ is finite, $\delta^a(t,\omega)$ may be a right-hand limit point of a certain $S_{\infty b}(\omega)$ (see Theorem 4.5), and not belong to $S_\infty^-(\omega)$.

It follows from the Corollary to Theorem 4.3 that if $0<s<t$, then

$$\mathbf{P}^a\{\delta_t^a\leqslant s;\ x_t^a=j\}=\sum_{i\in\mathbf{I}^a}\xi_i^a(s)f_{ij}(t-s)\quad (j\in\mathbf{I}^a); \tag{5.5}$$

$$\mathbf{P}^a\{\delta_t^a\leqslant s\}=\sum_{i\in\mathbf{I}^a}\xi_i^a(s)\,[1-L_i(t-s)]=1-\sum_{i\in\mathbf{I}^a}\xi_i^a(s)\,L_i(t-s). \tag{5.6}$$

In dealing with x^a, the appropriate state space is $\mathbf{I}^a$ (which may or may not include θ) as noted above, but we shall frequently omit it when no confusion can arise. We now define

$$\zeta_j^a(t)=\mathbf{P}^a\{\delta_t^a=0;\ x_t^a=j\} \tag{5.7}$$

$$\varrho^a(t)=\mathbf{P}^a\{\delta_t^a>0\}. \tag{5.8}$$

THEOREM 5.2. *We have*

$$\zeta_j^a(t)=\lim_{s\downarrow 0}\sum_i\xi_i^a(s)f_{ij}(t-s)=\lim_{s\downarrow 0}\sum_i\xi_i^a(s)f_{ij}(t); \tag{5.9}$$

$$\varrho^a(t)=\lim_{s\downarrow 0}\sum_i\xi_i^a(s)L_i(t-s)=\lim_{s\downarrow 0}\sum_i\xi_i^a(s)L_i(t); \tag{5.10}$$

$$\varrho^a(t)+\sum_j\zeta_j^a(t)=1; \tag{5.11}$$

$$\sum_i\zeta_i^a(s)\,f_{ij}(t)=\zeta_j^a(s+t). \tag{5.12}$$

Proof. The first equations in (5.9) and (5.10) follow upon letting $s \downarrow 0$ in (5.5) and (5.6); indeed in either case there is monotone convergence. The second equations follow from the stochastic continuity of δ_t^a and x_t^a. Analytically, the second equation in (5.9), e.g., can be proved by the inequalities, valid for $0 < s < t$, $0 < s < \delta$:

$$f_{ij}(t-s)\, f_{jj}(s) \leqslant f_{ij}(t) \leqslant f_{ij}(t+\delta-s)\,[f_{jj}(\delta-s)]^{-1},$$

and the continuity of ζ_j^a to be noted below. Summing (5.7) over j and adding (5.8) we obtain (5.11). Finally, (5.12) is obvious from the meaning of $\zeta_j^a(t)$ as the probability that $x^a(t, \omega) = j$ and $\omega \in 0^a(0, t)$. An analytic proof using (5.9) is also immediate if we observe that

$$\sum_i \xi_i^a(s)\, f_{ij}(t-s) \leqslant \sum_i \xi_i^a(s)\, p_{ij}(t-s) = \xi_j^a(t),$$

so that the series
$$\sum_j \{\sum_i \xi_i^a(s)\, f_{ij}(t-s)\} \tag{5.13}$$

in j is uniformly convergent in $s \in (0, t)$.

THEOREM 5.3. *For every a and b in* A, *there exists a nonnegative nondecreasing function $L^{ab} \leqslant 1$ and a sequence $s_n \downarrow 0$ such that*

$$L^{ab}(t) = \lim_{n\to\infty} \sum_i \xi_i^a(s_n)\, L_i^b(t-s_n) = \lim_{n\to\infty} \sum_i \xi_i^a(s_n)\, L_i^b(t) \tag{5.14}$$

for every $t > 0$. The function L^{ab} is absolutely continuous in $\mathbf{T}^0$ *but may have a jump at zero; its almost everywhere derivative l^{ab} satisfies, for almost every t, the equation*

$$l^{ab}(t) = \sum_j \zeta_j^a(s)\, l_j^b(t-s) \quad (0 < s < t). \tag{5.15}$$

Proof. The set of functions $\sum_i \xi_i^a(s) L_i^b(t-s)$ of t in (m^{-1}, ∞) indexed by $s \in (0, m^{-1})$ consists of bounded, nondecreasing functions. Hence by Helly's theorem a sequence $\{s_n\}$ exists for which the first relation in (5.14) holds for $t \in (m^{-1}, \infty)$. Letting $m \to \infty$ and using the diagonal argument we obtain the first relation in (5.14) as asserted. Now we have by (4.3), if $s < t_1 < t_2$;

$$\sum_i \xi_i^a(s)\, \{L_i^b(t_2-s) - L_i^b(t_1-s) = \}\sum_j \{\sum_i \xi_i^a(s)\, f_{ij}(t_1-s)\}\, L_j^b(t_2-t_1). \tag{5.16}$$

Letting $s \downarrow 0$ along the sequence $\{s_n\}$ and using (5.9), (5.14) and the remark involving (5.13), we obtain

$$L^{ab}(t_2) - L^{ab}(t_1) = \sum_j \zeta_j^a(t_1)\, L_j^b(t_2-t_1).$$

Consequently L^{ab} is continuous in $\mathbf{T}^0$ and the second relation in (5.14) follows from the first as in Theorem 5.2. Furthermore it follows that

$$L^{ab}(t_2) - L^{ab}(t_1) = \int_0^{t_2-t_1} \sum_j \zeta_j^a(t_1) l_j^b(u)\,du. \tag{5.17}$$

If $0 < r < s$, we have by (5.12) and (4.4):

$$\begin{aligned}\sum_j \zeta_j^a(s) l_j^b(u) &= \sum_j \{\sum_i \zeta_i^a(r) f_{ij}(s-r)\} l_j^b(u) \\ &= \sum_i \zeta_i^a(r) \sum_j f_{ij}(s-r) l_j^b(u) \\ &= \sum_i \zeta_i^a(r) l_i^b(u+s-r).\end{aligned} \tag{5.18}$$

Hence we can define a function l^{ab} by (5.15) and substituting into (5.17) we obtain

$$L^{ab}(t_2) - L^{ab}(t_1) = \int_{t_1}^{t_2} l^{ab}(s)\,ds \quad (0 < t_1 < t_2 < \infty). \tag{5.19}$$

COROLLARY, *The left member of (5.16) converges as $s \downarrow 0$ to the left member of (5.17).*

Proof. This follows since each sequence $\{s_n\}$ converging to zero contains a subsequence along which the left member of (5.16) converges to the unique limit given in (5.19).

Remark. The corollary says that the measures in t corresponding to $\sum_i \xi_i^a(s) L_i^b(t-s)$ converge on $\mathbf{T}^0$ as $s \downarrow 0$. We do not know if they converge on $\mathbf{T}$. More precisely, define $L^{ab}(0) = 0$ and let the jump of L^{ab} at zero be denoted by $L^{ab}(0+)$; does

$$\lim_{t\downarrow 0} \lim_{s\downarrow 0} \sum_i \xi_i^a(s) L_i^b(t-s)$$

exist and equal to $L^{ab}(0+)$?

THEOREM 5.4. *Under Assumptions A and B, we have*

$$\xi_j^a(t) = \zeta_j^a(t) + \sum_{b\in\mathbf{A}} \int_0^t \xi_j^b(t-s)\,dL^{ab}(s) \tag{5.20}$$

if and only if

$$\varrho^a(\cdot) = \sum_b L^{ab}(\cdot). \tag{5.21}$$

Proof. Substituting (5.1) into (4.21), we have

$$\xi_j^a(s+t) = \sum_i \xi_i^a(s) f_{ij}(t) + \sum_{b\in\mathbf{A}} \int_0^t \xi_j^b(t-u)\,d_u[\sum_i \xi_i^a(s) L_i^b(u)]. \tag{5.22}$$

Since A is denumerable, there exists by the diagonal argument a sequenee $\{s_n\} \downarrow 0$ such that (5.14) holds for every a and b. Since each ξ_j^b is continuous, this implies the convergence of each integral in (5.22) to its corresponding limit. Hence by (5.9) and Fatou's lemma, (5.20) holds with "$=$" replaced by "$\geqslant$". Summing over j in I^a, we have

$$1 \geqslant \sum_j \zeta_j^a(t) + \sum_{b \in \mathrm{A}} L^{ab}(t).$$

Comparing this with (5.11) we see that (5.21) is a necessary and sufficient condition for the equality to hold in (5.20).

COROLLARY. *A sufficient condition for the validity of (5.20) is that for each a, there is only a finite number of b such that $L^{ab}(\infty) > 0$; this is the case if* A *is a finite set.*

Let the Laplace transforms be defined as follows, $0 < \lambda < \infty$:

$$\left.\begin{aligned} \hat{\xi}_j^a(\lambda) &= \int_0^\infty e^{-\lambda t}\xi_j^a(t)\,dt, \quad \hat{\zeta}_j^a(\lambda) = \int_0^\infty e^{-\lambda t}\zeta_j^a(t)\,dt, \\ \hat{L}^{ab}(\lambda) &= \int_0^\infty e^{-\lambda t}dL^{ab}(t) = L^{ab}(0+) + \int_0^\infty e^{-\lambda t}l^{ab}(t)\,dt. \end{aligned}\right\} \tag{5.23}$$

The equation (5.20) becomes, omitting the index j:

$$\hat{\xi}^a(\lambda) = \hat{\zeta}^a(\lambda) + \sum_b \hat{L}^{ab}(\lambda)\,\hat{\xi}^b(\lambda); \tag{5.24}$$

or in matrix form on the super-index:

$$[I - \Lambda(\lambda)]\,\hat{\xi}(\lambda) = \hat{\zeta}(\lambda), \tag{5.25}$$

where I is the identity matrix,

$$\Lambda(\lambda) = (\hat{L}^{ab}(\lambda)), \quad (a, b) \in \mathrm{A} \times \mathrm{A},$$

and $\hat{\xi}(\lambda)$ and $\hat{\zeta}(\lambda)$ are regarded as column vectors with the components indexed by A.

Following an established terminology [1; § I. 3], we write $a \rightsquigarrow b$ iff $L^{ab}(\infty) > 0$, otherwise $a \not\rightsquigarrow b$; and we write $a \leftrightsquigarrow b$ iff $a \rightsquigarrow b$ and $b \rightsquigarrow a$. Note that an equivalent definition is obtained if we use $\hat{L}^{ab}(\lambda)$ for any $\lambda < \infty$ instead of $L^{ab}(\infty) = \hat{L}^{ab}(0)$. The index a is *essential* iff there exists b such that $a \rightsquigarrow b$ and for each such b we have also $b \rightsquigarrow a$; otherwise it is *inessential*. The relation $\leftrightsquigarrow$ among essential indices is an equivalence relation by means of which they are partitioned into *essential classes* $C_1, C_2, \ldots$. The set of inessential indices will be denoted by C_0.

THEOREM 5.5 *Suppose that* $\mathbf{A}$ *is a finite set. One of the following alternatives must occur:*

(i) *There exists an essential class* C *of indices such that*

$$\sum_{b\in C} L^{ab}(0+) = 1 \quad (a \in C). \tag{5.26}$$

In this case if $a \in C$ *and* $b \in C$, *then* ∞^a *and* ∞^b *are indistinguishable.* C *may be a singleton.*

(ii) *The matrix* $I - \Lambda(\lambda)$ *is invertible and*

$$\hat{\xi}(\lambda) = [I - \Lambda(\lambda)]^{-1} \hat{\zeta}(\lambda) = \sum_{n=0}^{\infty} \Lambda^n(\lambda) \hat{\zeta}(\lambda). \tag{5.27}$$

Proof. Suppose there exists an essential class C such that the restriction of $\Lambda(\lambda)$ to it is stochastic, namely, we have

$$\sum_{b\in C} \hat{L}^{ab}(\lambda) = 1 \quad (a \in C). \tag{5.28}$$

This can happen only if $\hat{L}^{ab}(\lambda) = L^{ab}(0+)$ and (5.26) holds. It follows from this, (5.10) and (5.14) that

$$\varrho^a(t) = \lim_{n\to\infty} \sum_i \xi_i^a(s_n) \sum_{b\in\mathbf{A}} L_i^b(t) = \sum_{b\in\mathbf{A}} L^{ab}(0+) = 1;$$

consequently $\varrho^a(0+) = \lim_{t\downarrow 0} \varrho^a(t) = 1$ and so $\zeta_j^a(t) = 0$ for every j and t by (5.11). Thus (5.24) reduces to

$$\hat{\xi}^a(\lambda) = \sum_{b\in C} L^{ab}(0+) \hat{\xi}^b(\lambda) \quad (a \in C). \tag{5.29}$$

Since C is a finite set, being a subset of $\mathbf{A}$, a well-known result in discrete parameter Markov chains asserts that the only solution $\hat{\xi}(\lambda)$ of such a system of equations is a constant. A simple algebraic proof is also available. Thus for each λ, $\hat{\xi}^a(\lambda) = \hat{\xi}^b(\lambda)$ for every a and b in C and we conclude that $\xi_j^a(t) \equiv \xi_j^b(t)$ identically in j and t for a and b in C, by the uniqueness of Laplace transforms.(1)

On the other hand, if there does not exist any essential class C with the property (5.28), then it is a consequence of the recurrence properties of discrete parameter Markov chains that the series $\sum_{n=0}^{\infty} \Lambda^n(\lambda)$ converges for $0 < \lambda < \infty$ and yields the inverse of $I - \Lambda(\lambda)$. Applying it to (5.25) we obtain (5.27), completing the proof of Theorem 5.5.

---

(1) This result should be compared with Theorem 4.6.

If indistinguishable boundary points have been merged, then the alternative (i) reduces to

$$L^{aa}(0+)=1$$

for a certain merged boundary point. This means τ^a is a left-hand limit point of $S_{\infty^*}(\omega)$ almost everywhere on Δ^a, and the fictitious state ∞^a behaves as an instantaneous state. Analytically, the equation (5.24) reduces to the trivial identity $\hat{\xi}^a(\lambda)=\hat{\xi}^a(\lambda)$.

Under the alternative (ii), the sample functions can be described as follows. For almost every ω, $x(\cdot,\omega)$ reaches the passable part of the boundary in a sequence of times

$$\tau(0)<\tau(1)<\tau(2)<\cdots \text{ (1)},$$

where $\tau(0)=\tau$ is the first infinity in our previous notation and if $\tau(n)=\infty$ then $\tau(n+1)=\tau(n+2)=\cdots=\infty$. We define

$$y(n)=\lim_{t\uparrow\tau(n)} x(t) \quad \text{if} \quad \tau(n)<\infty,$$

$$y(n)=\theta'' \qquad \text{if} \quad \tau(n)=\infty;$$

where the limit is taken in the metric topology of $\mathbf{J}^*$ so that $y(n)\in \mathbf{B}_0$. If we write simply "a" for "∞^a", the process $\{y(n), n\in \mathbf{N}\}$ is a discrete parameter homogeneous Markov chain with $\mathbf{A}_{\theta''}$ as its state space, and the stochastic completion (by θ'') of $(L^{ab}(\infty))$, $(a,b)\in \mathbf{A}\times\mathbf{A}$, as its one-step transition matrix. Furthermore we have

$$\mathbf{P}\{\tau(n+1)-\tau(n)\leqslant t;\, y(n+1)=b \mid y(n)=a\}=L^{ab}(t).$$

If we define $$z(t-\tau(0))=y(n) \quad \text{for} \quad \tau(n)\leqslant t<\tau(n+1),$$

the process $\{z(t), t\in \mathbf{T}\}$ is a so-called semi-Markovian process. Finally if we set

$$\nu(t)=n \quad \text{for} \quad \tau(n)\leqslant t<\tau(n+1),$$

then for any t and $\Lambda\in\mathfrak{F}_{\tau(\nu(t))}$,

$$\mathbf{P}\{x(t)=j \mid \Lambda;\, y(\nu(t))=a\}=\zeta_j^a(t-\tau(\nu(t))).$$

Perhaps it is better to describe the above situation in somewhat less precise but more intelligible terms as follows. The sample function of the Markov chain x is

(1) $\tau(n)$ is not the previous τ_n.

composed of a sequence of "waves" going from a passable atomic boundary point to another (not necessarily distinct one). The transition of these boundary points follows that of an imbedded Markov chain with $(L^{ab}(\infty))$ as one-step transition matrix. The length of each wave joining ∞^a to ∞^b has the distribution $L^{ab}(\cdot)$, independently of any occurrence outside this wave. Within each wave the sample function has only jumps, and consequently the transition of x there is by means of $\Phi=(f_{ij})$. If a wave begins at ∞^a, then at t units of time later x is in the state j with probability $\zeta_j^a(t)$. A sample function may have a finite number of waves before reaching a point on the recurrent part of the boundary and remaining there ever after; or it may have a final wave extending to infinity while approaching the impassable part of the boundary; or it may have an infinite sequence of waves going to infinity. Under the hypothesis of Theorem 5.5, in the case (ii), these waves cannot accumulate in the finite.

§ 6. The Second Approach

Let I' be the set of $\prod$-nonrecurrent states in $\mathbf{I}_\theta$; note that θ, if present, is recurrent.

For each a in $\mathbf{A}$ and j in $\mathbf{I}_\theta$, we set

$$g_j^a=\int_0^\infty \xi_j^a(t)\,dt. \tag{6.1}$$

Theorem 6.1. *The set of j for which $g_j^a>0$ is the state space $\mathbf{I}^a$ of the post-τ^a process. If $j\in\mathrm{I}^a$, then $g_j^a<\infty$ or $g_j^a=\infty$ according as $j\in\mathbf{I}'$ or $j\notin\mathrm{I}'$.*

Proof. The first assertion follows from the fact that I^a is the set of j for which $\xi_j^a(t)>0$ for some j and t, and the continuity of ξ_j^a. To prove the second assertion, we observe that if $j\in\mathrm{I}'$, then $\int_0^\infty p_{ij}(t)\,dt<\infty$. We have by (5.1)

$$p_{ij}(t)\geqslant\int_0^t \xi_j^a(t-s)\,dL_i^a(s);$$

consequently

$$\int_0^\infty p_{ij}(t)\,dt\geqslant L_i^a(\infty)\,g_j^a.$$

There exists an i in I such that $L_i^a(\infty)>0$; hence $g_j^a<\infty$. An alternative proof of this is as follows. If $S_j=S_j(\omega)=\{t\colon x(t,\omega)=j\}$ and μ is the Lebesgue measure on $\mathbf{T}$, then we have

$$g_j^a = \mathbf{E}\{\mu[S_j \cap (\tau^a, \infty)]\} \leqslant \mathbf{E}\{\mu[S_j]\} < \infty$$

by [1; Theorem II. 10.4].

On the other hand, if $j \notin \mathbf{I}'$ then $\int_0^\infty p_{jj}(t)\,dt = \infty$. We have by (4.21), for any s and t:

$$\xi_j^a(s+t) \geqslant \xi_j^a(s)\,p_{jj}(t).$$

It follows that for any u:

$$g_j^a \geqslant \int_0^u \xi_j^a(s+t)\,dt \geqslant \xi_j^a(s) \int_0^u p_{jj}(t)\,dt.$$

There exists s such that $\xi_j^a(s) > 0$; hence we obtain $g_j^a = \infty$ by letting $u \to \infty$.

Assumption C. $\mathbf{I} = \mathbf{I}'$; *namely there is no* $\prod$*-recurrent state except* θ *if present.*

It is not true that Assumption C can be made without loss of generality, even if we are only interested in the nonrecurrent part of the boundary. This is because a $\prod$-recurrent state need not be Φ-recurrent; see Theorem 3.2 and the Remark after it. In particular the Doob type of construction (see [1; Theorem II. 19.4]) leads to Π-recurrent states if $L_i(\infty) = 1$ for every i.

From now on in this section an unspecified index i, j or k is an element of $\mathbf{I}'$, and an unspecified sum over it is extended to $\mathbf{I}'$.

For $j \in \mathbf{I}'$, we set
$$\mathrm{H}_j^a(t) = g_j^a - \sum_i g_i^a f_{ij}(t). \tag{6.2}$$

This is the "dual" of (4.11).

Theorem 6.2. *We have if* $j \in \mathbf{I}^a$:

$$\mathrm{H}_j^a > 0, \quad \mathrm{H}_j^a \uparrow, \quad \mathrm{H}_j^a(0) = 0; \tag{6.3}$$

$$\mathrm{H}_j^a(s+t) - \mathrm{H}_j^a(t) = \sum_i \mathrm{H}_i^a(s)\,f_{ij}(t). \tag{6.4}$$

H_j^a *has a continuous positive derivative* η_j^a *in* $\mathbf{T}$ *satisfying*

$$\eta_j^a(s+t) = \sum_i \eta_i^a(s)\,f_{ij}(t). \tag{6.5}$$

Proof. We have

$$\sum_i g_i^a f_{ij}(t) = \sum_i \int_0^\infty \xi_i^a(s) f_{ij}(t)\, ds \leqslant \sum_i \int_0^\infty \xi_i^a(s)\, p_{ij}(t)\, ds$$
$$= \int_0^\infty \xi_j^a(s+t)\, ds = g_j^a - \int_0^t \xi_j^a(s)\, ds \leqslant g_j^a.$$

By the Appendix the second inequality above is strict, hence $H_j^a > 0$. Next we have

$$\sum_i H_i^a(s) f_{ij}(t) = \sum_i \{g_i^a - \sum_k g_k^a f_{ki}(s)\} f_{ij}(t) = \sum_i g_i^a f_{ij}(t) - \sum_k g_k^a f_{kj}(s+t) = H_j^a(s+t) - H_j^a(t);$$

hence $H_j^a \uparrow$. The continuous differentiability of H_j^a together with (6.5) follows from the equation (6.4) by a general lemma already cited under (4.3) in its dual form. Finally, $\eta_j^a > 0$ by the Appendix.

For every a and b in A, we set

$$\sigma^{ab}(t) = \sum_i g_i^a l_i^b(t). \tag{6.6}$$

It will follow from the proof below that the series in (6.6) converges for every $t > 0$, and is a nonincreasing function of t.

The next theorem is fundamental; it takes the place of Theorem 5.4 in the new approach.

THEOREM 6.3. *Under Assumptions A, B, and C, we have*

$$\int_0^t \eta_j^a(s)\, ds = \int_0^t \xi_j^a(s)\, ds + \sum_{b \in A} \int_0^t \sigma^{ab}(s)\, \xi_j^b(t-s)\, ds. \tag{6.7}$$

Proof. Let us rewrite (5.22) as

$$\xi_j^a(s+t) = \sum_i \xi_i^a(s) f_{ij}(t) + \sum_b \int_0^t [\sum_i \xi_i^a(s)\, l_i^b(u)]\, \xi_j^b(t-u)\, du. \tag{6.8}$$

For each b, there exist j and t_0 such that $\xi_j^b(t_0) > 0$, and this implies $\xi_j^b(t) > 0$ for all $t > t_0$ by (4.21) (for a stronger result see the Appendix). It follows from this that the series in square brackets in (6.8) converges for each fixed s and almost every u. Furthermore we can integrate to obtain

$$\infty > \int_0^\infty \xi_j^a(s+t)\, ds = \sum_i g_i^a f_{ij}(t) + \sum_b \int_0^t \sigma^{ab}(u)\, \xi_j^b(t-u)\, du, \tag{6.9}$$

where σ^{ab} is defined by (6.6) and the series there converges for almost every t. If it converges for t and $t < u$, then by (4.4),

$$\sum_i g_i^a l_i^b(u) = \sum_j [\sum_i g_i^a f_{ij}(u-t)]\, l_j^b(t) \leqslant \sum_j g_j^a l_j^b(t).$$

Hence the series in (6.6) converges for every $t > 0$. Finally, since

$$\int_0^\infty \xi_j^a(s+t)\,ds - \sum_i g_i^a f_{ij}(t) = g_j^a - \int_0^t \xi_j^a(s)\,ds - \sum_i g_i^a f_{ij}(t)$$

$$= \mathrm{H}_j^a(t) - \int_0^t \xi_j^a(s)\,ds$$

$$= \int_0^t \eta_j^a(s)\,ds - \int_0^t \xi_j^a(s)\,ds$$

by Theorem 6.2, (6.9) is equivalent to (6.7).

COROLLARY 1. $\mathrm{H}_j^a(\infty) = g_j^a$, *viz.*

$$\int_0^\infty \xi_j^a(s)\,ds = \int_0^\infty \eta_j^a(s)\,ds.$$

Proof. We have by (6.7) and (6.2)

$$\mathrm{H}_j^a(\infty) \geqslant \int_0^\infty \xi_j^a(s)\,ds = g_j^a \geqslant \mathrm{H}_j^a(t);$$

Corollary 1 follows upon letting $t \to \infty$.

COROLLARY 2. *σ^{ab} is summable over every finite interval.*

Proof. This follows from (6.7) and the fact, already used in the proof of the theorem, that for each b there exist j and t_0 such that $\xi_j^b(t) > 0$ for $t > t_0$, and consequently ξ_j^b is bounded away from zero in (t_0, t_1) for every $t_1 > t_0$, since it is continuous.

THEOREM 6.4. *For almost every t, the series*

$$\sum_i \eta_i^a(s)\, l_i^b(t-s) \tag{6.10}$$

converges for $0 < s < t$ and defines a function $\theta^{ab}(t)$ which does not depend on s. We have

$$\sigma^{ab}(t) = \int_t^\infty \theta^{ab}(s)\,ds = \sum_i \eta_i^a(t)\, L_i^b(\infty); \tag{6.11}$$

in particular, σ^{ab} is continuous in T^0.

Proof. If the series in (6.10) converges, then the sum does not depend on s for s in $(0, t)$ by exactly the same calculation as given in (5.18). Now by (6.6) and Corollary 1 to Theorem 6.3, we have

$$\sigma^{ab}(t) = \sum_i \int_0^\infty \eta_i^a(s)\, l_i^b(t)\,ds = \int_0^\infty \theta^{ab}(s+t)\,ds \tag{6.12}$$

for almost every t, using Fubini's theorem on product measures; similarly

$$\sigma^{ab}(t) = \int_0^\infty \sum_i \eta_i^a(t) l_i^b(s)\, ds = \sum_i \eta_i^a(t) L_i^b(\infty). \tag{6.13}$$

Since both the extreme members of (6.12) are nonincreasing and the one on the right is continuous, (6.12) must hold for every $t > 0$. Now the first member of (6.13) is nonincreasing and continuous, while the last one is easily seen to be nonincreasing, hence (6.13) must hold for every t. Theorem 6.4 is proved.

To proceed further we need an essential strengthening of Assumption B, already imposed in Theorem 5.5.

ASSUMPTION B′. A *is a finite set.*

Let us put
$$\sigma^a(t) \equiv \sum_{b \in \mathbf{A}} \sigma^{ab}(t); \tag{6.14}$$

$$\eta_*^a(t) \equiv \sum_i \eta_i^a(t). \tag{6.15}$$

THEOREM 6.5. *Under Assumption B′ the function η_*^a is finite, nonincreasing and continuous in $\mathbf{T}^0$, and summable in every finite interval. We have*

$$\eta_*^a(t) - \eta_*^a(\infty) = \sigma^a(t); \tag{6.16}$$

$$\eta_*^a(\infty) = \sum_i \eta_i^a(t)\, [1 - L_i(\infty)]. \tag{6.17}$$

Proof. Summing (6.7) over j and using Corollary 2 to Theorem 6.3, we have

$$\sum_j \mathrm{H}_j^a(t) \leqslant t + \sum_b \int_0^t \sigma^{ab}(s)\, ds < \infty, \tag{6.18}$$

since b ranges over a finite set. It follows that the series in (6.15) converges for almost every t. If $\eta_*^a(t) < \infty$ then we have by (6.5), for every $t' \geqslant 0$:

$$\eta_*^a(t) = \sum_i \eta_i^a(t) \{\sum_j f_{ij}(t') + L_i(t')\} = \eta_*^a(t+t') + \sum_i \eta_i^a(t) L_i(t'). \tag{6.19}$$

Hence $\eta_*^a(t+t') < \infty$ and we conclude that η_*^a is finite and nonincreasing in $\mathbf{T}^0$. Its continuity there also follows from (6.19), since each L_i is continuous in $\mathbf{T}$. The summability of η_*^a follows from (6.18). Finally, rewriting (4.3) as

$$1 - L_i(s+t) = \sum_j f_{ij}(s)\, [1 - L_j(t)]$$

and letting $t \to \infty$ we see that

$$1 - L_i(\infty) = \sum_j f_{ij}(s)\,[1 - L_j(\infty)]. \tag{6.20}$$

On the other hand, summing (6.13) over b we have

$$\sigma^a(t) = \sum_i \eta_i^a(t)\, L_i(\infty). \tag{6.21}$$

It follows from (6.20) and (6.21) that

$$\begin{aligned}\eta_*^a(t) - \sigma^a(t) &= \sum_j \eta_j^a(t)\,[1 - L_j(\infty)]\\ &= \sum_j \sum_i \eta_i^a(s)\, f_{ij}(t-s)\,[1 - L_j(\infty)]\\ &= \sum_i \eta_i^a(s)\,[1 - L_i(\infty)] = \eta_*^a(s) - \sigma^a(s).\end{aligned} \tag{6.22}$$

Thus $\eta_*^a(t) - \sigma^a(t)$ is a constant which must be $\eta_*^a(\infty)$ since $\sigma^a(\infty) = 0$.

Corollary. $\eta_*^a(\infty) \leqslant 1$.

Proof. Divide (6.18) by t and let $t \to \infty$.

The next two theorems are valid under Assumptions A, B and C (without B′). Remember that $j \neq \theta$ below.

Theorem 6.7. *η_j^a is absolutely continuous and*

$$\eta_j^a(t) = -\sum_i g_i^a f'_{ij}(t), \tag{6.23}$$

where the series converges absolutely for $t \geqslant 0$; in particular

$$\eta_j^a(0) = -\sum_i g_i^a q_{ij}. \tag{6.24}$$

We have for almost every t:

$$\frac{d}{dt}\eta_j^a(t) = \sum_i \eta_i^a(t)\, q_{ij}. \tag{6.25}$$

Proof. We have

$$0 \leqslant \frac{\mathrm{H}_j^a(t)}{t} = \sum_i g_i^a \frac{[\delta_{ij} - f_{ij}(t)]}{t}$$

or

$$\sum_{i \neq j} g_i^a \frac{f_{ij}(t)}{t} \leqslant g_j^a \frac{1 - f_{jj}(t)}{t}.$$

Letting $t \downarrow 0$ we have by Fatou's lemma

$$\sum_{i \neq j} g_i^a q_{ij} \leqslant g_j^a q_j. \tag{6.26}$$

Next, we have, using the second system of equations (II) in § 2 for Φ:

$$\sum_i g_i^a |f_{ij}'(t)| \leqslant \sum_i g_i^a \{f_{ij}(t) q_j + \sum_{k \neq j} f_{ik}(t) q_{kj}\}$$
$$\leqslant \{\sum_i g_i^a f_{ij}(t)\} g_j + \sum_{k \neq j} \{\sum_i g_i^a f_{ik}(t)\} q_{kj}$$
$$\leqslant g_j^a q_j + \sum_{k \neq j} g_k^a q_{kj} \leqslant 2 g_j^a q_j \tag{6.27}$$

by (6.26). It follows that

$$\mathrm{H}_j^a(t) = \sum_i g_i^a [\delta_{ij} - f_{ij}(t)] = -\sum_i g_i^a \int_0^t f_{ij}'(s)\, ds = -\int_0^t \sum_i g_i^a f_{ij}'(s)\, ds \tag{6.28}$$

by (6.27) and bounded convergence. Upon differentiation we obtain (6.23). Starting with (6.23), substituting from (II) again,. and relying on (6.27) for the interchange of summations, we obtain

$$\eta_j^a(t) = -\sum_i g_i^a \{\sum_k f_{ik}(t) q_{kj}\} = -\sum_k \{\sum_i g_i^a f_{ik}(t)\} q_{kj} = -\sum_k \{g_k^a - \mathrm{H}_k^a(t)\} q_{kj}$$
$$= -\sum_k \int_t^\infty \eta_k^a(s) q_{kj}\, ds. \tag{6.29}$$

Letting $t \downarrow 0$ we obtain (6.24) by Corollary 1 to Theorem 6.3. Furthermore, the series in (6.25) converges absolutely, having only one negative term, for almost every t and the summation and integration in the last member of (6.29) can be interchanged, proving the absolute continuity of η_j^a together with (6.25).

THEOREM 6.8. *ξ_j^a is absolutely continuous in* $\mathbf{T}$; *we have*

$$\frac{d}{dt} \xi_j^a(t) + \xi_j^a(t) q_j = \sum_i \xi_i^a(s) v_{ij}(t-s) \tag{6.30}$$

for almost every t and every s in $(0, t)$, *where*

$$v_{ij}(t) = p_{ij}'(t) + p_{ij}(t) q_j; \text{(1)} \tag{6.31}$$

$$\frac{d}{dt} \xi_j^a(t) \geqslant \sum_i \xi_i^a(t) q_{ij} \tag{6.32}$$

for almost every t. For each a in A *and j in* $\mathbf{I}^a$, *the following three propositions are equivalent:*

(i) (II_{ij}) *holds for every i in* $\mathbf{I}^a$;

(ii) *Equality holds in* (6.32) *for almost every t;*

(iii) $\xi_j^a(0) = \eta_j^a(0) = 0$.

(1) See [1; § II. 16] for a discussion of v_{ij}.

Proof. Using [1; (II.16.2)] we write for each $t>0$ and $0<s<t$,

$$\xi_j^a(t)=\sum_i \xi_i^a(s)\,p_{ij}(t-s)=\sum_i \xi_i^a(s)\left[\delta_{ij}e^{-q_j(t-s)}+\int_0^{t-s} v_{ij}(u)\,e^{-q_j(t-s-u)}du\right]. \qquad (6.33)$$

Since $\sum_i \xi_i^a(s)=1$ by (4.20), this shows that ξ_j^a is absolutely continuous in $\mathbf{T}^0$, hence in $\mathbf{T}$ by its continuity at zero. Multiplying (6.33) through by $e^{q_j t}$ and using Fubini's theorem on differentiation, we obtain (6.30) for almost every t. Substituting the inequality

$$v_{ij}(t-s)\geqslant \sum_{k\neq j} p_{ik}(t-s)\,q_{kj} \qquad (6.34)$$

into (6.30) and using (4.21) we obtain (6.32).

If (i) is true, then equality holds in (6.34) by the definition of v_{ij} and the preceding substitution leads to equality in (6.32). Thus (i) implies (ii). Conversely if (ii) is true, then for almost every t and $0<s<t$,

$$\frac{d}{dt}\xi_j^a(t)+\xi_j^a(t)\,q_j=\sum_{k\neq j}\xi_k^a(t)\,q_{kj}=\sum_{k\neq j}\Big[\sum_i \xi_i^a(s)\,p_{ik}(t-s)\Big]q_{kj}=\sum_i \xi_i^a(s)\sum_{k\neq j}p_{ik}(t-s)\,q_{kj}. \qquad (6.35)$$

Comparing (6.30) with (6.35), we see that equality must hold in (6.34) whenever $\xi_i^a(s)>0$. For each i in $\mathbf{I}^a$ this is the case for every sufficiently large s. It follows that equality holds for every i in $\mathbf{I}^a$ and every $t-s$, that is, (II_{ij}) holds for every i in $\mathbf{I}^a$. Thus (ii) implies (i) and we have proved the equivalence of (i) and (ii).

It follows from (6.32) that for each $u>0$:

$$\xi_j^a(u)-\xi_j^a(0)-\sum_i\int_0^u \xi_i^a(t)\,q_{ij}\,dt=\int_0^u\left[\frac{d}{dt}\xi_j^a(t)-\sum_i \xi_i^a(t)\,q_{ij}\right]dt\geqslant 0. \qquad (6.36)$$

Since $\lim_{u\to\infty}\xi_j^a(u)=0$ as a consequence of Theorem 6.1, we have upon letting $u\to\infty$ and using (6.24):

$$-\xi_j^a(0)+\eta_j^a(0)=-\xi_j^a(0)-\sum_i g_i^a\,q_{ij}\geqslant 0. \qquad (6.37)$$

If (iii) is true, then there is equality in (6.36) and hence also in (6.32). Thus (iii) implies (ii). Conversely if (ii) is true, the same argument shows that there is equality in (6.37). To prove that $\xi_j^a(0)=0$, let us suppose the contrary. There exist i and t such that $L_i^a(t)>0$, hence we have, as a consequence of the Corollary to Theorem 4.3 and Theorem 4.4:

$$\mathbf{P}_i\{\tau^a\leqslant t;\ x_s=j\ \text{ for all }\ s\ \text{ in }\ (\tau^a,t)\}\geqslant\int_0^t \xi_j^a(0)\,e^{-q_j(t-s)}\,dL_i^a(s)>0.$$

This means that there is positive probability that $x_0=i$, $x_t=j$ and that the last dis-

continuity of the sample function before time t is a "pseudo-jump" from ∞^a and not a jump. By [1; Theorem II.14.4] (¹) this cannot happen under (i). Since (i) and (ii) have been shown to be equivalent, we conclude that (ii) implies (iii) and Theorem 6.8 is completely proved.

Remark. If (6.7) holds, then dividing through by t and letting $t \downarrow 0$ we obtain $\eta_j^a(0) \geqslant \xi_j^a(0)$ in general, and $\eta_j^a(0) = \xi_j^a(0)$ if $\sigma^{ab}(0) < \infty$ for every a and b.

§ 7. The Dual Chain

In this section we study the notion of a dual chain. Combining it and the results of § 6 we shall derive a representation of $\{\xi_j\}$ when the method of § 5 fails, and discuss the case left open there, namely the alternative (i) in Theorem 5.5. This is the case where the boundary behavior, even under the most stringent set of assumptions made here, is still not fully understood. It should be stressed that the dual chain studied here is more an analytical device than a genuinely probabilistic one. The latter would be that of a *reversed chain* as has been introduced in simpler cases (see [2]) and would involve an investigation of the sample function as the direction of time is reversed. This has not yet been done in a satisfactory manner and the results below serve only as a sort of vague reflection of the true state of matters.

For a few moments Assumptions A, B, and C (without B′) will suffice. For each a in A we set

$$\tilde{p}_{ji}^a(t) = \frac{g_i^a p_{ij}(t)}{g_j^a}. \tag{7.1}$$

The matrix $(\tilde{p}_{ji}^a)$, $(j, i) \in \mathrm{I}^a \times \mathrm{I}^a$, will be called the *a-dual to* (p_{ij}). Where this dual matrix is concerned the index set will be I^a without specific mention.

THEOREM 7.1. *For each a in A, $(\tilde{p}_{ji}^a)$ isa strictly substochastic transition matrix. Its initial derivative matrix $(\tilde{q}_{ji}^a)$ and the corresponding minimal solution $(\tilde{f}_{ji}^a)$ are given as follows:*

$$\tilde{q}_{ji}^a = \frac{g_i^a q_{ij}}{g_j^a}, \tag{7.2}$$

$$\tilde{f}_{ji}^a(t) = \frac{g_i^a f_{ij}(t)}{g_j^a}. \tag{7.3}$$

(¹) I take this opportunity to acknowledge that the argument in the first few lines of p. 223 in [1] is inadequate for an instantaneous state k, as pointed out to me by S. Orey. This has been corrected both by him and by myself but the revision is too long to be included here. For the purpose here, where all states are stable, the proof given in [1] is correct.

The matrix $(\tilde{q}^a_{ji})$ *is stochastic if and only if* $\eta^a_j(0)=0$ *for every* j *in* $\mathbf{I}^a$. *We have*

$$\tilde{L}^a_j(t) = 1 - \sum_i \tilde{f}^a_{ji}(t) = \frac{\mathrm{H}^a_j(t)}{g^a_j}; \tag{7.4}$$

the function $\tilde{L}^a_j$ *has a continuous derivative* $\tilde{l}^a_j$ *satisfying*

$$\sum_i [g^a_i \tilde{l}^a_i(t)] f_{ij}(s) = g^a_j \tilde{l}^a_j(t+s). \tag{7.5}$$

Proof. It is easy to see that $(\tilde{p}^a_{ji})$ satisfies the semi-group property corresponding to (2.2). Next we have by definition

$$\sum_i \tilde{p}^a_{ji}(t) = \frac{1}{g^a_j} \sum_i \int_0^\infty \xi^a_i(s) p_{ij}(t)\, ds = \frac{1}{g^a_j} \int_t^\infty \xi^a_j(s)\, ds \leq 1. \tag{7.6}$$

In fact if $j \in \mathbf{I}^a$, then $\xi^a_j(\cdot) > 0$ by the Appendix, so that there is strict inequality in (7.6) for every $t > 0$. This and trivial inspection show that $(\tilde{p}^a_{ji})$ is a strictly sub-stochastic standard transition matrix and (7.2) follows at once from (7.1). It is easy to verify that $(\tilde{f}^a_{ij})$ as defined by (7.3) is the minimal solution to the two systems of Kolmogorov differential equations (I) and (II) in § 2 when (q_{ij}) there is replaced by $(\tilde{q}^a_{ij})$. Moreover we have

$$\sum_i \tilde{q}^a_{ji} = \frac{1}{g^a_j} \sum_i g^a_i q_{ij} = -\frac{\eta^a_j(0)}{g^a_j} \leq 0 \tag{7.7}$$

by (6.24). The equation (7.4) follows at once from (7.3) and (6.2). Hence by Theorem 6.2, we have

$$\tilde{l}^a_j(t) = \frac{\eta^a_j(t)}{g^a_j} \tag{7.8}$$

and (7.5) follows from (6.5). Theorem 7.1 is proved.

A homogeneous Markow chain $\tilde{x}^a = \{\tilde{x}^a(t), t \in \mathrm{T}\}$ having $\mathbf{I}^a$ as its state space and the stochastic completion of $(\tilde{p}^a_{ji})$ as its transition matrix is called an *a-dual* to x. If $\eta^a_j(0) = 0$ for every j in $\mathbf{I}^a$, then it satisfies an assumption corresponding to Assumption A for x and so we may proceed to apply the preceding theory to it. However, to encompass as large a state space as possible we must take a suitable mixture of the indices a as follows.

Let $$\tilde{I} = \bigcup_{a \in \mathrm{A}} \mathbf{I}^a.$$

$\tilde{I}$ is the state space of the post-τ process under Assumptions A and B; it is clear that it is a $\prod$-stochastically closed subset of $\mathbf{I}$. Let

$$L^a = \sum_i \gamma_i L_i^a,$$

where γ is the initial distribution of x, and set

$$h_j = \sum_{a \in \mathbf{A}} L^a(\infty) g_j^a. \tag{7.9}$$

We now introduce the following assumption which is essentially the dual of Assumption A.

Assumption $\tilde{\text{A}}$. *The second system of Kolmogorov differential equations holds.*

Theorem 7.2. (*Under Assumptions* A, B, C *and* $\tilde{\text{A}}$.) *We have* $h_j < \infty$ *for every* j *in* $\mathbf{I}$, *and* $h_j > 0$ *if and only if* $j \in \tilde{\mathbf{I}}$. *Furthermore,*

$$\sum_i h_i p_{ij}(t) \leqslant h_j, \tag{7.10}$$

$$\sum_i h_i q_{ij} = 0. \tag{7.11}$$

Proof. The first assertion follows from (5.1) upon integration over $\mathbf{T}$:

$$h_j = \int_0^\infty \sum_i \gamma_i [p_{ij}(t) - f_{ij}(t)]\, dt < \infty,$$

since $\int_0^\infty \sum_i \gamma_i p_{ij}(t)\, dt < \infty$ for a nonrecurrent state j. By the definition of A, for each a in A there exists an i such that $\gamma_i > 0$ and $L_i^a(\infty) > 0$, hence $L^a(\infty) > 0$. By Theorem 6.1, $g_j^a > 0$ if and only if $j \in \mathbf{I}^a$. These remarks prove that $h_j > 0$ if and only if $j \in \tilde{\mathbf{I}}$. Next, we have

$$\sum_i h_i p_{ij}(t) = \sum_a L^a(\infty) \sum_i g_i^a p_{ij}(t) \leqslant \sum_a L^a(\infty) g_j^a = h_j.$$

Finally, we have by (6.24) and Theorem 6.8, under Assumption $\tilde{\text{A}}$:

$$\sum_i h_i q_{ij} = \sum_a L^a(\infty) \sum_i g_i^a q_{ij} = 0,$$

the interchange of the repeated summation being justified since $\sum_i |h_i q_{ij}| \leqslant 2 h_j q_j < \infty$ by (6.26). Theorem 7.2 is proved.

We now set

$$\tilde{p}_{ji}(t) = \frac{h_i p_{ij}(t)}{h_j}, \tag{7.12}$$

$$\tilde{q}_{ji} = \frac{h_i q_{ij}}{h_j}, \quad -\tilde{q}_{jj} = \tilde{q}_j = q_j, \tag{7.13}$$

$$\tilde{r}_{ji} = \frac{(1-\delta_{ji})\tilde{q}_{ji}}{\tilde{q}_j} = \frac{(1-\delta_{ij})\, h_i q_{ij}}{h_j q_j}, \tag{7.14}$$

$$\tilde{f}_{ji}(t) = \frac{h_i f_{ij}(t)}{h_j}. \tag{7.15}$$

The matrix $\tilde{\Pi} = (\tilde{p}_{ji})$, $(j, i) \in \tilde{\mathbf{I}} \times \tilde{\mathbf{I}}$, will be called *the dual transition matrix* to (p_{ij}); similarly for $\tilde{Q}$, $\tilde{\mathrm{P}}$ and $\tilde{\Phi}$. A homogeneous Markov chain $\tilde{x} = \{\tilde{x}(t),\ t \in \mathbf{T}\}$ having $\tilde{\mathbf{I}}$ as its state space and $\tilde{\Pi}_{\tilde{\theta}}$ as its transition matrix will be called the *dual chain* to x. By virtue of (7.11), the matrices $\tilde{Q}$ and $\tilde{\mathrm{P}}$ are stochastic and so the dual chain satisfies the assumption corresponding to Assumption A and we can define its jump chain $\tilde{\chi} = \{\tilde{\chi}_n\}$, its Martin boundary $\tilde{\mathbf{B}}$ and the passable part $\tilde{\mathbf{B}}$. The assumption corresponding to Assumption B, which we now make, is as follows.

ASSUMPTIONS $\tilde{\mathrm{B}}$, $\tilde{\mathrm{B}}'$. *The passable part of the dual boundary is completely atomic.*

These atoms will be denoted by $\{\tilde{\infty}^{\tilde{a}},\ \tilde{a} \in \tilde{\mathrm{A}}\}$. Assumption $\tilde{\mathrm{B}}$ becomes Assumption $\tilde{\mathrm{B}}'$ iff $\tilde{\mathrm{A}}$ is a finite set. Under Assumption $\tilde{\mathrm{B}}'$ we may and shall replace the definition (7.9) by the simpler one:

$$h_j = \sum_a g_j^a.$$

It is clear that Theorem 7.2 remains valid after this replacement.

THEOREM 7.3. *Under Assumptions* $\tilde{\mathrm{A}}$ *and* $\tilde{\mathrm{B}}$*:*

$$p_{ij}(t) = f_{ij}(t) + \sum_{\tilde{a} \in \tilde{\mathrm{A}}} \int_0^t k_i^{\tilde{a}}(s)\, \psi_j^{\tilde{a}}(t-s)\, ds, \tag{7.16}$$

where

$$\sum_j p_{ij}(s)\, k_j^{\tilde{a}}(t) = k_i^{\tilde{a}}(s+t), \tag{7.17}$$

$$\sum_i \psi_i^{\tilde{a}}(s)\, f_{ij}(t) = \psi_j^{\tilde{a}}(s+t). \tag{7.18}$$

Proof. Theorem 5.1 applied to the dual transition matrix yields:

$$\tilde{p}_{ji}(t) = \tilde{f}_{ji}(t) + \sum_{\tilde{a} \in \tilde{\mathrm{A}}} \int_0^t \tilde{l}_j^{\tilde{a}}(s)\, \tilde{\xi}_i^{\tilde{a}}(t-s)\, ds, \tag{7.19}$$

where

$$\sum_j \tilde{f}_{ij}(t)\, \tilde{l}_j^{\tilde{a}}(s) = \tilde{l}_i^{\tilde{a}}(t+s), \tag{7.20}$$

$$\sum_i \tilde{\xi}_i^{\tilde{a}}(t)\, \tilde{p}_{ij}(s) = \tilde{\xi}_j^{\tilde{a}}(t+s), \tag{7.21}$$

these formulas being the duals of (4.4) and (4.21) respectively. Putting

$$h_j l_j^{\tilde{a}}(\cdot) = \psi_j^{\tilde{a}}(\cdot), \tag{7.22}$$

$$h_i^{-1} \tilde{\xi}_i^{\tilde{a}}(\cdot) = k_i^{\tilde{a}}(\cdot), \tag{7.23}$$

and substituting from (7.12) and (7.15) we obtain (7.17) and (7.18). Theorem 7.3 is proved.

Clearly the coexistence of the two formulas (5.1) and (7.16) has interesting implications. However, due to evident technical difficulties more stringent assumptions than those needed for both formulas will be invoked in the next theorem. We must also introduce a new definition.

A passable atomic boundary point ∞^a is called *nonrepeatable* iff for every $i \in \tilde{\mathrm{I}}$ we have $L_i^a(\infty) = 0$; otherwise it is called *repeatable*. Let the subset of A corresponding to nonrepeatable boundary points be A_0. Such a boundary point ∞^a is reached exactly once on Δ^a, and is never reached again after the first infinity. It is inessential according to the definition in § 5, indeed $L^{ba} \equiv 0$ for every b in A. It is trivial to construct a nonrepeatable boundary point: we need only start the Markov chain with an ascending escalator and hitch on an open Markov chain, say a descending escalator, with a disjoint state space.

THEOREM 7.4. *Under Assumptions* A, B′, $\tilde{\mathrm{A}}$, $\tilde{\mathrm{B}}'$ *and* C, *there exist nondecreasing, bounded functions* $M^{a\tilde{a}}$, $(a, \tilde{a}) \in \mathrm{A} \times \tilde{\mathrm{A}}$, *such that for every* a *in* A, j *in* I *and* t *in* T *we have*

$$\xi_j^a(t) = \sum_{\tilde{a} \in \tilde{\mathrm{A}}} \int_0^t \psi_j^{\tilde{a}}(t-s)\, dM^{a\tilde{a}}(s). \tag{7.24}$$

Proof. Comparing (5.1) and (7.16) we have

$$\sum_{b \in \mathrm{A}} \int_0^t l_i^b(s)\, \xi_j^b(t-s)\, ds = \sum_{\tilde{b} \in \tilde{\mathrm{A}}} \int_0^t k_i^{\tilde{b}}(s) \psi_j^{\tilde{b}}(t-s)\, ds, \tag{7.25}$$

where both A and $\tilde{\mathrm{A}}$ are finite sets. According to the Corollary to Theorem 4.1, for each a there exists a sequence $\{i_n\}$ such that for every b in A,

$$\lim_{n \to \infty} \int_0^t l_{i_n}^b(s)\, \xi_j^b(t-s)\, ds = \delta^{ba} \xi_j^a(t). \tag{7.26}$$

If $a \notin \mathrm{A}_0$, namely if ∞^a is repeatable, then we may choose $\{i_n\}$ so that $i_n \in \tilde{I}$ for every n and (7.25) holds with $i = i_n$. Now if we integrate this equation over T and take only the term corresponding to the index $\tilde{a}$ on the right side, we have by (7.22) and (7.23):

$$\int_0^\infty k_i^{\tilde a}(s)\,ds \int_0^\infty \psi_j^{\tilde a}(s)\,ds = \int_0^\infty k_i^{\tilde a}(s)\,ds\, h_j \tilde L_j^{\tilde a}(\infty) \leqslant \sum_{b\in A} L_i^b(\infty) g_j^b \leqslant h_j. \tag{7.27}$$

Choosing j such that $\tilde L_j^{\tilde a}(\infty) > 0$, and putting $K_i^{\tilde a}(t) = \int_0^t k_i^{\tilde a}(s)\,ds$, we see that

$$K_i^{\tilde a}(\infty) \leqslant \frac{1}{\tilde L_j^{\tilde a}(\infty)} < \infty. \tag{7.28}$$

Thus the family of nondecreasing functions $\{K_i^{\tilde a}, i \in \mathrm{I}\}$ has a uniformly bounded total variation and so is weakly compact. It follows that there is a subsequence $\{i_n'\}$ (depending on $\tilde a$) of $\{i_n\}$ (depending on a) for which $K_{i_n'}^{\tilde a}(\cdot)$ converges weakly to a limit $M^{a\tilde a}(\cdot)$ which is nondecreasing and bounded with $M^{a\tilde a}(0) = 0$. Applying this result to (7.25), noting the continuity of $\psi_j^{\tilde a}$ and using (7.26), we obtain (7.24) for every $a \notin \mathrm{A}_0$.

It remains to prove (7.24) for $a \in \mathrm{A}_0$. Since $L^{ab}(\infty) = 0$ for $a \in \mathrm{A}$ and $b \in \mathrm{A}_0$, we may rewrite (5.20) as follows, omitting the index j:

$$\xi^a(t) = \zeta^a(t) + \sum_{b\in \mathrm{A}\setminus \mathrm{A}_0} \int_0^t \xi^b(t-s)\,dL^{ab}(s). \tag{7.29}$$

Substituting from the proved part of (7.24), we have

$$\xi^a(t) = \zeta^a(t) + \sum_{\tilde a\in\tilde{\mathrm{A}}} \int_0^t \psi^{\tilde a}(t-s)\,dN^{a\tilde a}(s) \tag{7.30}$$

where

$$N^{a\tilde a} = \sum_{b\in \mathrm{A}\setminus \mathrm{A}_0} (L^{ab} * M^{b\tilde a}), \tag{7.31}$$

and $*$ denotes a convolution. Next, the equation (5.12) becomes, after substituting from (7.15) and noting that $\zeta_i^a(\cdot) \equiv 0$ if $i \notin \tilde{\mathrm{I}}$,

$$\sum_{i\in\tilde{\mathrm{I}}} \tilde f_{ji}(t)[h_i^{-1}\zeta_i^a(s)] = h_j^{-1}\zeta_j^a(t+s) \quad (j\in\tilde{\mathrm{I}}).$$

Using the definition given in § 4, the set $\{h_i^{-1}\zeta_i^a(\cdot)\}$ is an exit solution for $\tilde\Phi$ for each a, since $\zeta_i^a(0) = 0$ by (7.29) and Theorem 6.8; moreover by (5.20),

$$\int_0^\infty h_i^{-1}\zeta_i^a(t)\,dt \leqslant h_i^{-1}\int_0^\infty \xi_i^a(t)\,dt \leqslant 1.$$

Hence according to Theorem 4.2 applied to the dual chain, and using (7.22):

$$\zeta_i^a(t) = \sum_{\tilde a\in\tilde{\mathrm{A}}} c^{a\tilde a}\psi_i^{\tilde a}(t) \tag{7.32}$$

where $0 \leqslant c^{a\tilde{a}} \leqslant 1$. Substituting into (7.30), we see that (7.24) holds with

$$M^{a\tilde{a}} = N^{a\tilde{a}} + c^{a\tilde{a}}\,\varepsilon.$$

Theorem 7.4 is proved.

A set of nonnegative functions $\{u_i(\cdot)\}$ with $u_i(0) = 0$ for every i and satisfying the system of functional equations

$$\sum_i u_i(s)\,f_{ij}(t) = u_j(s+t) \tag{7.33}$$

will be called an *entrance solution* for Φ. Under Assumption $\tilde{A}$, the sets $\{\zeta_i^a(\cdot)\}$ and $\{\eta_i^a(\cdot)\}$ defined in Theorems 5.2 and 6.2 are entrance solutions for Φ. We have seen in the above how an entrance solution for Φ corresponds to an exit solution for $\tilde{\Phi}$. The set $\{\psi^{\tilde{a}},\ \tilde{a} \in \tilde{A}\}$ forms an extreme base for the space of entrance solutions restricted to $\tilde{I}$. In particular, we can express ζ and η in terms of ψ.

Corollary to Theorem 7.4. *We have for every a and $j \in I$;*

$$\zeta_j^a(t) = \sum_{\tilde{a}} M^{a\tilde{a}}(0+)\,\psi_j^{\tilde{a}}(t), \tag{7.34}$$

$$\eta_j^a(t) = \sum_{\tilde{a}} M^{a\tilde{a}}(\infty)\,\psi_j^{\tilde{a}}(t). \tag{7.35}$$

Proof. (7.34) follows from the following calculation:

$$\zeta_j^a(t) = \lim_{s \downarrow 0} \sum_i \sum_{\tilde{a}} \int_0^s \psi_i^{\tilde{a}}(s-u)\,f_{ij}(t-s)\,dM^{a\tilde{a}}(u)$$

$$= \sum_{\tilde{a}} \lim_{s \downarrow 0} \int_0^s \psi_j^{\tilde{a}}(t-u)\,dM^{a\tilde{a}}(u) = \sum_{\tilde{a}} M^{a\tilde{a}}(0+)\,\psi_j^a(t).$$

To prove (7.35), we first integrate (7.24) to obtain

$$g_j^a = \sum_{\tilde{a}} M^{a\tilde{a}}(\infty) \int_0^\infty \psi_j^{\tilde{a}}(s)\,ds.$$

Consequently
$$\sum_i g_i^a\, f_{ij}(t) = \sum_{\tilde{a}} M^{a\tilde{a}}(\infty) \int_0^\infty \psi_j^{\tilde{a}}(s+t)\,ds,$$

and
$$H_j^a(t) = \sum_{\tilde{a}} M^{a\tilde{a}}(\infty) \int_0^t \psi_j^{\tilde{a}}(s)\,ds,$$

from which (7.35) follows upon differentiation.

On the basis of Theorem 7.4, we can express the probabilities in (5.5) and (5.6) in a suggestive way as follows:

$$\mathbf{P}^a\{\delta_t^a \leqslant s\} = \sum_{\tilde{a}} \int_0^s \psi_*^{\tilde{a}}(t-u)\, dM^{a\tilde{a}}(u),$$

$$\mathbf{P}^a\{\delta_t^a \leqslant s;\ x_t^a = j\} = \sum_{\tilde{a}} \int_0^s \psi_j^{\tilde{a}}(t-u)\, dM^{a\tilde{a}}(u)$$

$$= \sum_{\tilde{a}} \int_0^s \frac{\psi_j^{\tilde{a}}(t-u)}{\psi_*^{\tilde{a}}(t-u)}\, \psi_*^{\tilde{a}}(t-u)\, dM^{a\tilde{a}}(u).$$

Thus the last non-jump discontinuity before time t in the post-τ^a process enjoys properties similar to that of the last exit time from an ordinary state before time t, discussed in [2]. Starting from this it is possible to discuss the reversed chain rigorously as a probabilistic object. but we shall not pursue the matter further here.

We can also use Theorem 7.4 to obtain criteria for either alternative in Theorem 5.5. Write $\psi_*^{\tilde{a}}(\cdot) = \sum_j \psi_j^{\tilde{a}}(\cdot)$ as in (6.15).

THEOREM 7.5. *Under Assumptions* A *and* B′, *if* $\eta_*^a(0) < \infty$ (*or equivalently* $\sigma^a(0) < \infty$) *then* $\varrho^a(0) = 0$. *Case* (ii) *of Theorem 5.5 obtains if in each essential class of indices, there exists at least one index* a *for which* $\eta_*^a(0) < \infty$. *Under the additional Assumptions* $\tilde{\mathrm{A}}$ *and* $\tilde{\mathrm{B}}$ *this is the case if* $\psi_*^{\tilde{a}}(0) < \infty$ *for every* $\tilde{a}$. *On the other hand, if* $\psi_*^{\tilde{a}}(0) = \infty$ *for every* $\tilde{a}$, *then case* (i) *of Theorem 5.5 obtains.*

Proof. It follows from (5.10) that

$$\varrho^a(t) = \lim_{u \downarrow 0} u^{-1} \int_0^u \sum_i \xi_i^a(s) L_i(t-s)\, ds.$$

By (6.7), we have $\int_0^u \xi_i^a(s)\, ds \leqslant \int_0^u \eta_i^a(s)\, ds$, hence by (6.19):

$$\varrho^a(t) \leqslant \lim_{u \downarrow 0} u^{-1} \int_0^u [\eta_*^a(s) - \eta_*^a(t)]\, ds = \eta_*^a(0) - \eta_*^a(t)$$

since $\eta_*^a(s)$ is nondecreasing as $s \downarrow 0$. Hence if $\eta_*^a(0) < \infty$, then

$$\varrho^a(0) = \lim_{t \downarrow 0} \varrho^a(t) \leqslant \lim_{t \downarrow 0} [\eta_*^a(0) - \eta_*^a(t)] = 0.$$

Since A is finite, (5.21) holds and

$$\sum_b L^{ab}(0+) = 0. \tag{7.36}$$

If a is essential this excludes the possibility of (5.26) for the essential class to which a belongs. If this is so for each such class, Theorem 5.5 asserts that case (ii) there occurs. Finally, if $\psi_*^{\tilde{a}}(0)<\infty$ for every $\tilde{a}$, then $\eta_*^a(0)<\infty$ by (7.35) and we have case (ii) by the above.

Now suppose the other extreme: $\psi_*^{\tilde{a}}(0)=\infty$ for every $\tilde{a}$. As before we have

$$\sum_i \psi_i^{\tilde{a}}(s)L_i(t-s)=\psi_*^{\tilde{a}}(s)-\psi_*^{\tilde{a}}(t),$$

and consequently by (7.24):

$$\sum_i \xi_i^a(s)L_i(t-s)=\sum_{\tilde{a}}\int_0^s [\psi_*^{\tilde{a}}(s-u)-\psi_*^{\tilde{a}}(t-u)]\,dM^{a\tilde{a}}(u).$$

It follows that
$$\sum_{\tilde{a}} M^{a\tilde{a}}(0+)[\psi_*^{\tilde{a}}(s)-\psi_*^{\tilde{a}}(t)]\leqslant 1.$$

As $s\downarrow 0$ this implies $M^{a\tilde{a}}(0+)=0$ for every $\tilde{a}$ and so by (7.34), $\zeta_j^a(t)=0$ for every j and t. Hence by (5.11), $\varrho^a(t)=1$ for every t and so $\varrho^a(0)=1$. This means $\sum_t L^{ab}(0+)=1$ by (5.21) and we have case (i) of Theorem 5.3.

§ 8. The Construction Theorem

There is a basic connection between Theorems 6.3 and 7.4 which leads to a solution of the construction problem. In this section we make full use of the method of Laplace transforms.

Taking Laplace transforms in (6.7) and using matrix notation, we have

$$\hat{\eta}(\lambda)=[I+\lambda\hat{\Sigma}(\lambda)]\hat{\xi}(\lambda), \tag{8.1}$$

where $\hat{\Sigma}(\lambda)$ is the matrix $(\hat{\sigma}^{ab}(\lambda))$, $(a,b)\in\mathrm{A}\times\mathrm{A}$. We are under Assumptions B′ and $\tilde{\mathrm{B}}'$ so that both A and $\tilde{\mathrm{A}}$ are finite sets. We have by (7.35),

$$\hat{\eta}(\lambda)=M\hat{\psi}(\lambda), \tag{8.2}$$

where $M=(M^{a\tilde{a}}(\infty))$, $(a,\tilde{a})\in\mathrm{A}\times\tilde{\mathrm{A}}$, is a constant matrix. For a few moments let $\theta_0^{\tilde{a}b}$ and $\sigma_0^{\tilde{a}b}$ denote the quantities θ^{ab} and σ^{ab} in (6.10) and (6.11) when η^a is replaced by $\psi^{\tilde{a}}$. Since both $\{\eta_i^a(\cdot)\}$ and $\{\psi_i^{\tilde{a}}(\cdot)\}$ are entrance solutions, the properties of θ^{ab} and σ^{ab} deriving from the fact that η^a is an entrance solution hold also for $\theta_0^{\tilde{a}b}$ and $\sigma_0^{\tilde{a}b}$. Finally, let

$$u^{\tilde{a}a}(\lambda)=\lambda\hat{\sigma}_0^{\tilde{a}a}(\lambda) \tag{8.3}$$

and $U(\lambda)$ be the matrix $(u^{\tilde{a}a}(\lambda))$, $(\tilde{a},a)\in\tilde{\mathrm{A}}\times\mathrm{A}$.

LEMMA. *For each* $\tilde{a}$ *and* a, $\dfrac{d}{d\lambda}\, u^{\tilde{a}a}(\lambda)$ *is a completely monotonic function of* λ.

Proof. We have by (6.12) and a simple calculation:

$$\frac{d}{d\lambda}u^{\tilde{a}a}(\lambda)=\int_0^\infty e^{-\lambda t}(1-\lambda t)\,\sigma_0^{\tilde{a}a}(t)\,dt$$

$$=\int_0^\infty e^{-\lambda t}(1-\lambda t)\,dt\int_t^\infty \theta_0^{\tilde{a}a}(s)\,ds=\int_0^\infty e^{-\lambda s}s\theta_0^{\tilde{a}a}(s)\,ds. \tag{8.4}$$

Since we have
$$\theta_0^{\tilde{a}a}(s)=\frac{1}{s}\int_0^s \sum_i \psi_i^{\tilde{a}}(u)\,l_i^a(s-u)\,du$$

by Theorem 6.4 applied to ψ, the last member of (8.4) is equal to $\sum_i \hat{\psi}_i^{\tilde{a}}(\lambda)\,\hat{l}_i^a(\lambda)$. Since $\hat{l}_i^a(\lambda)\leqslant 1$, and $\sum_i \psi_i^{\tilde{a}}(\lambda)<\infty$ by Theorem 6.5 applied to ψ, $\sum_i \hat{\psi}_i^{\tilde{a}}(\lambda)\,\hat{l}_i^a(\lambda)$ converges and is completely monotonic in λ since each term is. The lemma is proved.

In terms of ψ, the equation (8.1) can be written as

$$M\hat{\psi}(\lambda)=[I+MU(\lambda)]\,\hat{\xi}(\lambda). \tag{8.5}$$

It is our object to study the solvability of (8.5) for $\hat{\xi}(\lambda)$. The following theorem is a general result about completely monotonic functions. Let us call a matrix of functions completely monotonic iff each element of the matrix is so.

THEOREM 8.1. *Let* M *be an* $\mathrm{A}\times\tilde{\mathrm{A}}$ *matrix with elements which are nonnegative constants,* $M(\lambda)$ *likewise with elements which are nonnegative functions of* λ; $U(\lambda)$ *an* $\tilde{\mathrm{A}}\times\mathrm{A}$ *matrix with elements whose derivatives are completely monotonic functions of* λ. *Suppose that for each* λ *we have*

$$M=[I+MU(\lambda)]\,M(\lambda), \tag{8.6}$$

where I *is the* $\mathrm{A}\times\mathrm{A}$ *identity matrix. Then both* $I+MU(\lambda)$ *and* $I+U(\lambda)\,M$ *are invertible; we have*

$$M=M(\lambda)\,[I+U(\lambda)\,M]; \tag{8.7}$$

and the matrix $M(\lambda)$ *is completely monotonic.*

Proof. To show that $I+MU(\lambda)$ is invertible, suppose there exists a vector v such that

$$v[I+MU(\lambda)]=0. \tag{8.8}$$

Then by (8.6), $vM=0$ and consequently by (8.8), $v=0$. To show that $I+U(\lambda)M$ is invertible, suppose there exists a vector w such that

$$[I+U(\lambda)M]w=0. \tag{8.9}$$

Then $$[I+MU(\lambda)]Mw=M[I+U(\lambda)M]w=0.$$

Since $I+MU(\lambda)$ is invertible as just shown, we have $Mw=0$ and consequently $w=0$ by (8.9).(1)

Since $$[I+MU(\lambda)]M=M[I+U(\lambda)M],$$

it follows that $$M[I+U(\lambda)M]^{-1}=[I+MU(\lambda)]^{-1}M=M(\lambda)$$

by (8.6), and so (8.7) is true. Finally, for $\delta>0$ consider

$$\begin{aligned}&[I+MU(\lambda+\delta)]\{M(\lambda+\delta)-M(\lambda)\}[I+U(\lambda)M]\\ &\quad=[I+MU(\lambda+\delta)]\{[I+MU(\lambda+\delta)]^{-1}M-M[I+U(\lambda)M]^{-1}\}[I+U(\lambda)M]\\ &\quad=M[I+U(\lambda)M]-[I+MU(\lambda+\delta)]M=M[U(\lambda)-U(\lambda+\delta)]M.\end{aligned}$$

Dividing through by δ and letting $\delta\downarrow 0$, we obtain

$$[I+MU(\lambda)]M'(\lambda)[I+U(\lambda)M]=-MU'(\lambda)M.$$

Equivalently, by (8.6) and (8.7), we have

$$-M'(\lambda)=M(\lambda)U'(\lambda)M(\lambda). \tag{8.10}$$

For the sake of induction let us now suppose that

$$(-1)^m M^{(m)}(\lambda)\geqslant 0 \quad (0\leqslant m\leqslant n). \tag{8.11}$$

This is true for $m=0$ by hypothesis. Differentiating (8.10) n times by Leibniz's rule, we have

$$\begin{aligned}(-1)^{n+1}M^{(n+1)}(\lambda)&=\sum_{0\leqslant j+k\leqslant n}\frac{n!}{j!\,k!\,(n-j-k)!}M^{(j)}(\lambda)\,U^{(n+1-j-k)}(\lambda)\,M^{(k)}(\lambda)\\ &=\sum_{0\leqslant j+k\leqslant n}\frac{n!}{j!\,k!\,(n-j-k)!}(-1)^j M^{(j)}(\lambda)(-1)^{n-j-k}\\ &\qquad\times U^{(n+1-j-k)}(-1)^k M^k(\lambda)\geqslant 0,\end{aligned}$$

(1) I am indebted to N. G. de Bruijn for the preceding proof.

by the induction hypothesis and the hypothesis about $U(\lambda)$. Therefore (8.11) is true also for $m=n+1$ and the induction is complete, proving that $M(\lambda)$ is completely monotonic.

THEOREM 8.2. *There exists an* $\mathrm{A}\times\tilde{\mathrm{A}}$ *matrix* $M(\lambda)$ *such that*

$$\hat{\xi}(\lambda)=M(\lambda)\,\hat{\psi}(\lambda)\,(^1)\quad(0<\lambda<\infty),\tag{8.12}$$

if and only if there exists a constant matrix M *such that*

$$\hat{\eta}(\lambda)=M\hat{\psi}(\lambda)\quad(0<\lambda<\infty),\tag{8.13}$$

and such that $I+MU(\lambda)$ *is invertible. In this case* $M(\lambda)$ *is completely monotonic and* $M=M(0)$.

Proof. Suppose (8.12) holds, namely

$$\hat{\xi}_i^a(\lambda)=\sum_{\tilde{a}} m^{a\tilde{a}}(\lambda)\,\hat{\psi}_i^{\tilde{a}}(\lambda).\tag{8.14}$$

By (7.22) and the Corollary to Theorem 4.1, for each $\tilde{a}\in\tilde{\mathrm{A}}$ there exists a sequence $\{i_n\}$ in $\tilde{\mathrm{I}}$ such that

$$\lim_{n\to\infty} h_{i_n}^{-1}\,\hat{\psi}_{i_n}^{\tilde{b}}(\lambda)=\delta^{\tilde{a}\tilde{b}}\quad(\tilde{b}\in\tilde{\mathrm{A}}).\tag{8.15}$$

It follows from this and (8.14) that

$$m^{a\tilde{a}}(\lambda)=\lim_{n\to\infty} h_{i_n}^{-1}\,\hat{\xi}_{i_n}^a(\lambda)\geqslant 0,$$

so that the matrix $M(\lambda)$ is automatically nonnegative for every λ. Next, there exists a constant matrix M and a sequence $\{\lambda_n\}$ converging to zero such that $\lim_{n\to\infty} M(\lambda_n)=M$; each element $m^{a\tilde{a}}$ of M is finite since by (8.14):

$$m^{a\tilde{a}}\leqslant\frac{\hat{\xi}_i^a(0)}{\hat{\psi}_i^{\tilde{a}}(0)}=\frac{g_i^a}{h_i\tilde{L}_i^{\tilde{a}}(\infty)}<\infty.$$

It follows that
$$g_i^a=\sum_{\tilde{a}} m^{a\tilde{a}}\int_0^\infty \psi_i^{\tilde{a}}(s)\,ds;$$

and consequently as in the proof of (7.35) that

$$\eta_i^a(t)=\sum_{\tilde{a}} m^{a\tilde{a}}\,\psi_i^{\tilde{a}}(t).\tag{8.16}$$

(¹) The $M(\cdot)$ here is the Laplace transform of the $M(\cdot)$ in Theorem 7.4. We have omitted the cumbersome ^ where confusion is unlikely.

Taking Laplace transforms we obtain (8.13). Furthermore, substituting from (8.12) into (8.5), we have

$$M\hat{\psi}(\lambda) = [I + MU(\lambda)]\, M(\lambda)\, \hat{\psi}(\lambda).$$

Since the set $\{\psi^{\tilde{a}}(\lambda), \tilde{a} \in \tilde{\mathrm{A}}\}$ is linearly independent for each λ, a fact which is obvious from (8.15), it follows that

$$M = [I + MU(\lambda)]\, M(\lambda).$$

Theorem 8.1 is therefore applicable to yield the conclusions that $I + MU(\lambda)$ is invertible and that $M(\lambda)$ is completely monotonic.

Conversely, suppose that (8.13) holds; then $M \geqslant 0$ by (8.15). If $I + MU(\lambda)$ is invertible, then

$$\hat{\xi}(\lambda) = [I + MU(\lambda)]^{-1} M\hat{\psi}(\lambda),$$

and so if we set

$$M(\lambda) = [I + MU(\lambda)]^{-1} M, \tag{8.17}$$

we obtain (8.12) and $M(\lambda)$ is completely monotonic as before. Theorem 8.2 is proved.

COROLLARY 1. *$\lim_{\lambda \to 0} M(\lambda)$ exists.*

Proof. This follows from the uniqueness of the representation in (8.16).

We have formulated Theorem 8.2 in such a way as to stress the logical equivalence of two analytical propositions. Actually we know (8.12) is true under our assumptions by Theorem 7.4, hence the new fact that emerges is as follows.

COROLLARY 2. *The matrix $I + \lambda\Sigma(\lambda)$ in (8.1), or equivalently the matrix $I + MU(\lambda)$ in (8.5), is invertible.*

Let us recount the main steps of analysis up to this point. We are given a substochastic transition matrix $\prod$ on the index set $\mathrm{I} \times \mathrm{I}$ to begin with. The initial derivative matrix Q and the minimal solution Φ are then defined. Assumptions A, $\tilde{\mathrm{A}}$, B′ and C are made. We then define l, ξ and $\tilde{Q}$. Now Assumption $\tilde{\mathrm{B}}'$ is made, and ψ is defined. The following decomposition (or representation) formula ensues by virtue of Theorems 5.1 and 7.4:

$$\hat{\prod}(\lambda) = \hat{\Phi}(\lambda) + \hat{l}(\lambda)\, \hat{M}(\lambda)\, \hat{\psi}(\lambda), \tag{8.18}$$

where $\hat{\prod}(\lambda) = (\hat{p}_{ij}(\lambda))$, $\hat{\Phi}(\lambda) = (\hat{f}_{ij}(\lambda))$, $(i, j) \in \mathrm{I} \times \mathrm{I}$, and where $\hat{M}(\lambda)$ is written for the $M(\lambda)$ in (8.12) in conformity with the rest of our notation. $U(\lambda)$ is defined through l and ψ; finally let us write $\langle u, v\rangle = \sum_i u_i v_i$ if $u = \{u_i\}$ and $v = \{v_i\}$; $1 = \{1\}$; and set

$$M = \hat{M}(0) \quad (= M(\infty)),$$

$$\beta = \lim_{\lambda \to 0} \langle \lambda \hat{\psi}(\lambda), 1 \rangle \quad (= \psi_*(\infty)),$$

then we have:

$$I + MU(\lambda) \text{ and } I + U(\lambda) M \text{ are both invertible for every } \lambda; \tag{8.19}$$

$$\langle M\beta, 1 \rangle \leqslant 1. \tag{8.20}$$

The last inequality is equivalent to the Corollary to Theorem 6.6.

The full converse of the above will now be proved. We are given Q on $\mathbf{I}\times\mathbf{I}$ to begin with satisfying Assumption A, from which Φ is defined. Let $\{l^a, a \in \mathrm{A}\}$ be a finite set of exit solutions and $\{\psi^{\tilde{a}}, \tilde{a} \in \tilde{\mathrm{A}}\}$ a finite set of entrance solution for Φ; and define $U(\lambda)$ as in (8.3). Let M be an $\mathrm{A}\times\tilde{\mathrm{A}}$ matrix with nonnegative constant elements satisfying (8.19) and (8.20). Now define $\hat{\prod}(\lambda)$ by (8.18).

Theorem 8.3. *$\hat{\prod}(\lambda)$ is the Laplace transform of a substochastic transition matrix with Q as its initial derivative matrix; and every such $\hat{\prod}(\lambda)$ can be constructed in this way under Assumptions* A, B′, $\tilde{\mathrm{A}}$, $\tilde{\mathrm{B}}'$ *and* C.

Proof. We have already proved the second part of the theorem.

The calculations for the first part will be briefly indicated, omitting the ^ on the Laplace transforms. We shall first verify the resolvent equation for $\prod$:

$$\prod(\mu) - \prod(\lambda) = (\lambda - \mu) \prod(\lambda) \prod(\mu) \quad \left(0 < \frac{\lambda}{\mu} < \infty\right). \tag{8.21}$$

We being by writing down similar equations for Φ, l and ψ:

$$\Phi(\mu) - \Phi(\lambda) = (\lambda - \mu)\, \Phi(\lambda)\, \Phi(\mu), \tag{8.22}$$

$$l(\mu) - l(\lambda) = (\lambda - \mu)\, \Phi(\lambda)\, l(\mu), \tag{8.23}$$

$$\psi(\mu) - \psi(\lambda) = (\lambda - \mu)\, \psi(\lambda)\, \Phi(\mu), \tag{8.24}$$

the last two being the double Laplace transforms of (4.4) and (7.18). Next, we define $\theta^{\tilde{a}a}(t)$ (1) as in (6.10) to be $\sum_i \psi_i^{\tilde{a}}(s)\, l_i^a(t-s)$, $\hat{\theta}^{\tilde{a}a}(\lambda)$ to be its Laplace transform and $\Theta(\lambda) = (\hat{\theta}^{\tilde{a}a}(\lambda))$, $(\tilde{a}, a) \in \tilde{\mathrm{A}} \times \mathrm{A}$. Set also

$$\theta^{\tilde{a}a}(\lambda, \mu) = \sum_i \psi_i^{\tilde{a}}(\lambda)\, l_i^a(\mu) = \langle \psi^{\tilde{a}}(\lambda), l^a(\mu) \rangle,$$

$$\Theta(\lambda, \mu) = (\theta^{\tilde{a}a}(\lambda, \mu)) \quad ((\tilde{a}, a) \in \tilde{\mathrm{A}} \times \mathrm{A}).$$

(1) This is the $\theta_0^{\tilde{a}a}(t)$ used momentarily in the second paragraph of the section.

It follows from a computation based on Theorem 6.4 with η replaced by ψ that

$$\Theta(\mu)-\Theta(\lambda)=(\lambda-\mu)\,\Theta(\lambda,\mu).$$

Finally, by the relation corresponding to the first equation in (6.11), we have

$$U(\mu)-U(\lambda)=\Theta(\lambda)-\Theta(\mu). \tag{8.25}$$

Hence it follows from (8.17) that

$$\begin{aligned}&[I+MU(\lambda)]\,[M(\lambda)-M(\mu)]\,[I+U(\mu)\,M]\\ &\quad=M[I+U(\mu)\,M]-[I+MU(\lambda)]\,M=M[U(\mu)-U(\lambda)]\,M\\ &\quad=M[\Theta(\lambda)-\Theta(\mu)]\,M=(\mu-\lambda)\,M\Theta(\lambda,\mu)\,M\end{aligned}$$

or equivalently

$$(\lambda-\mu)\,M(\lambda)\,\Theta(\lambda,\mu)\,M(\mu)=M(\mu)-M(\lambda). \tag{8.26}$$

Now we have, upon substitution from (8.18):

$$\begin{aligned}\Pi(\lambda)\,\Pi(\mu)=\Phi(\lambda)\,\Phi(\mu)+l(\lambda)\,M(\lambda)\,\psi(\lambda)\,\Phi(\mu)+\Phi(\lambda)\,l(\mu)\,M(\mu)\,\psi(\mu)\\ +\,l(\lambda)\,M(\lambda)\,\Theta(\lambda,\mu)\,M(\mu)\,\psi(\mu).\end{aligned}$$

Hence using (8.22), (8.23), (8.24), and (8.26), we have

$$\begin{aligned}&(\lambda-\mu)\,\Pi(\lambda)\,\Pi(\mu)\\ &\quad=\Phi(\lambda)-\Phi(\mu)+l(\lambda)\,M(\lambda)\,[\psi(\mu)-\psi(\lambda)]+[l(\mu)-l(\lambda)]\,M(\mu)\,\psi(\mu)\\ &\qquad\qquad+l(\lambda)\,[M(\mu)-M(\lambda)]\,\psi(\mu)\\ &\quad=\Phi(\mu)+l(\mu)\,M(\mu)\,\psi(\mu)-\Phi(\lambda)-l(\lambda)\,M(\lambda)\,\psi(\lambda)\\ &\quad=\Pi(\mu)-\Pi(\lambda).\end{aligned}$$

Thus (8.21) is true. Next, we have by the relations corresponding to (6.16) with η replaced by ψ:

$$\langle\lambda\psi(\lambda),1\rangle=U(\lambda)\,1+\beta.$$

Hence it follows from (8.12), (8.17) and (8.20) that

$$\begin{aligned}\langle\lambda\xi(\lambda),1\rangle&=M(\lambda)\,[U(\lambda)\,1+\beta]\\ &=[I+MU(\lambda)]^{-1}[MU(\lambda)+M\beta]\\ &\leqslant[I+MU(\lambda)]^{-1}[MU(\lambda)+I]\,1=1,\end{aligned} \tag{8.27}$$

and consequently for each $i\in I$, if Π_i and Φ_i denote the ith rows of Π and Φ:

$$\begin{aligned}\langle\lambda\Pi_i(\lambda),1\rangle&=\langle\lambda\Phi_i(\lambda),1\rangle+\sum_a l_i^a(\lambda)\,\langle\lambda\xi^a(\lambda),1\rangle\\ &\leqslant 1-l_i(\lambda)+\sum_a l_i^a(\lambda)=1.\end{aligned}$$

Thus $\prod(\lambda)$ is substochastic, and is stochastic if, and only if, $M\beta = 1$. Theorem 8.3 is proved.

The condition (8.19) can be made more explicit in particular cases. The following theorem is due to Feller [7; Theorem 14.1].

THEOREM 8.4. *Suppose $U(\infty) < \infty$; then every $\hat{\prod}(\lambda)$ can be constructed in the following way. Choose an $A \times \tilde{A}$ matrix N with nonnegative constant elements satisfying the condition*

$$N[U(\infty)\,1 + \beta] \leqslant 1; \tag{8.28}$$

set

$$M(\lambda) = \{I - N[U(\infty) - U(\lambda)]\}^{-1} N, \tag{8.29}$$

and define $\hat{\prod}(\lambda)$ by (8.18).

Proof. We prove only the necessity of (8.28) and (8.29); their sufficiency can be verified as in the preceding proof. Let us rewrite (8.7) as

$$M(\lambda) = [I - M(\lambda)\,U(\lambda)]\,M. \tag{8.30}$$

Letting $\lambda \to \infty$ and writing N for $M(\infty)$, we have

$$N = [I - NU(\infty)]\,M. \tag{8.31}$$

It follows from the condition (8.28) that $NU(\infty)$ is substochastic, and consequently each row of $N[U(\infty) - U(\lambda)]$ has a sum which is strictly less than one since the vanishing of a row sum in $NU(\lambda)$ implies that of the corresponding row sum in $NU(\infty)$. Hence the matrix

$$I - N[U(\infty) - U(\lambda)] \tag{8.32}$$

is invertible. The preceding argument is taken from Feller [7]. Now we have by (8.31),

$$I - N[U(\infty) - U(\lambda)] = [I - NU(\infty)]\,[I + MU(\lambda)]. \tag{8.33}$$

By Corollary 2 to Theorem 8.2, the second factor on the right side of (8.33), as well as the product, is invertible. Hence the first factor is also invertible by elementary matrix theory. We conclude by (8.17), (8.31) and (8.33) that

$$\begin{aligned} M(\lambda) &= [I + MU(\lambda)]^{-1} M = [I + MU(\lambda)]^{-1}[I - NU(\infty)]^{-1} N \\ &= \{I - N[U(\infty) - U(\lambda)]\}^{-1} N. \end{aligned} \tag{8.34}$$

Next, we have from (8.31), $\quad M = [I - NU(\infty)]^{-1} N$

and consequently (8.20) becomes

$$N\beta \leqslant [I - NU(\infty)]\,1$$

which is (8.28).

Note that if we write properly $\hat{M}(\lambda)$ for the $M(\lambda)$ above and use $M(t)$ as in Theorem 7.4, we have $M = \hat{M}(0) = M(\infty)$ and $N = \hat{M}(\infty) = M(0+)$, and we infer that the two sets $\{\zeta^a\}$ and $\{\eta^a\}$, $a \in A$, in (7.34) and (7.35) are linear combinations of each other. Apart from this additional information, the proof of Theorem 8.4 is unnecessarily complicated. Indeed (8.29) is a special case of our earlier Theorem 5.5, as to be shown now. By Theorem 7.5, the hypothesis that $U(\infty) < \infty$ implies that case (ii) in Theorem 5.5 occurs. By (7.34), we have $\hat{\zeta}(\lambda) = N\hat{\psi}(\lambda)$ where N is as before. Using the notation in (5.27) and in the proof of Theorem 8.3, we have

$$\Lambda(\lambda) = N\Theta(\lambda) = N[U(\infty) - U(\lambda)],$$

the last equation being a consequence of (8.25). Substituting into (5.27), we obtain (8.29) by comparison with (8.12). Theorem 8.4 is proved.

The case where the matrix $U(\infty)$ contains infinite elements will now be sketched following Feller. A diagonal element of the matrix $I - M(\lambda)\,U(\lambda)$ is of the form

$$1 - \lambda \sum_i \hat{\xi}_i^a(\lambda)\,L_i^a(\infty) = \lambda \int_0^\infty e^{-\lambda t}[1 - \sum_i \xi_i^a(t)\,L_i^a(\infty)]\,dt,$$

hence positive unless $L_i^a(\infty) = 1$ for each $i \in \mathrm{I}^a$. It is easy to see that this is impossible under Assumption C. Hence we can write

$$I - M(\lambda)\,U(\lambda) = D(\lambda)\,[I - \bar{S}(\lambda)], \tag{8.35}$$

where $D(\lambda)$ is the diagonal part of the matrix on the left side of (8.35) and $\bar{S}(\lambda)$ has zero elements on the diagonal. Now define $\bar{M}(\lambda)$ by

$$\bar{M}(\lambda) = [D(\lambda)]^{-1}\,M(\lambda) = [I - \bar{S}(\lambda)]\,M, \tag{8.36}$$

where the second equation follows from (8.30). Letting $\lambda_n \to \infty$ so that

$$\lim_{n\to\infty} \bar{M}(\lambda_n) = \bar{M}, \quad \lim_{n\to\infty} \bar{S}(\lambda_n) = \bar{S}, \tag{8.37}$$

we obtain

$$\bar{M} = [I - \bar{S}]\,M. \tag{8.38}$$

Thus $\bar{M}$ and $\bar{S}$ take the place of N and $NU(\infty)$ respectively in (8.31). Substituting (8.6) into (8.38), we obtain

$$\bar{M} = [I - \bar{S}]\,[I + MU(\lambda)]\,M(\lambda) = [I - \bar{S} + \bar{M}U(\lambda)]\,M(\lambda). \tag{8.39}$$

Let us classify the indices in A according to the substochastic matrix $\bar{S}$, in a similar way as in § 5. Just as there, the matrix $I-\bar{S}$ is invertible unless there exists an essential class C such that for every a in C, the ath row in $\bar{S}$ has sum equal to one. Since $\bar{S}$ is zero on the diagonal, such a class C must contain more than one index. It follows from (8.38) that if $M_C, \bar{M}_C$ denote the restrictions of $M, \bar{M}$ to $C\times\tilde{\mathrm{A}}$, but $\bar{S}_C, (I-\bar{S})_C$ those of $\bar{S}, I-\bar{S}$ to $C\times C$, we have

$$M_C=\bar{M}_C+\bar{S}_C M_C \geqslant \bar{S}_C M_C. \tag{8.40}$$

Now a general theorem about discrete paramater Markov chains states that an excessive (superregular) function bounded below on a recurrent class is a constant. (In the case of a finite class as here, a simple algebraic proof is obtained by considering the minimum value of the function.) Applying this to (8.40) we infer that equality holds in (8.40) so that $\bar{M}_C=0$, and consequently we have by (8.39):

$$[I-\bar{S}]_C M_C(\lambda)=0.$$

It follows from (8.12) that, if $\hat{\xi}_C(\lambda)$ denotes the restriction of $\hat{\xi}(\lambda)$ to C:

$$[I-\bar{S}]_C\hat{\xi}_C(\lambda)=[I-\bar{S}]_C M_C(\lambda)\hat{\psi}(\lambda)=0,$$

and so
$$\hat{\xi}_C(\lambda)=\bar{S}_C\hat{\xi}_C(\lambda).$$

Applying again the theorem just cited, we see that $\hat{\xi}(\lambda)$ is constant on C. This being true for every λ, we conclude that $\xi^a(t)\equiv\xi^b(t)$ for every a and b in C. Thus the boundary points ∞^a for a in C are all indistinguishable from each other. If this eventuality is excluded, then $I-\bar{S}$ is invertible, and so is $I-\bar{S}+\bar{M}U(\lambda)$. We have therefore proved the following result.

THEOREM 8.5. *If all boundary points are distinguishable from each other, then we have*

$$M(\lambda)=[I-\bar{S}+\bar{M}U(\lambda)]^{-1}\bar{M}. \tag{8.41}$$

This was proved by Feller under the superfluous assumption that every element of $U(\lambda)$ be positive.(1) For the consequent construction theorem similar to Theorem 8.4 above, we refer to Feller [7].

(1) I am indebted to David Williams for a verification of Feller's theorem by a purely algebraic method, which leads to the disposition above.

§ 9. The One Exit Case

We make Assumptions A, B and C and the further assumption that the set A in Assumption B consists of one element only. The index a corresponding to this element will be omitted, thus e.g., $l_i(t) = l_i^a(t)$.

In this case we have

$$\sigma(t) = \sum_i g_i l_i(t) = \sum_i \eta_i(t) L_i(\infty),$$

$$u(\lambda) = \lambda\hat{\sigma}(\lambda) = \lambda \int_0^\infty e^{-\lambda t} \sigma(t)\,dt;$$

and (8.1) reduces to
$$\hat{\eta}(\lambda) = [1 + u(\lambda)]\,\hat{\xi}(\lambda),$$
or
$$\hat{\xi}(\lambda) = m(\lambda)\,\hat{\eta}(\lambda), \tag{9.1}$$
where
$$m(\lambda) = \frac{1}{1 + u(\lambda)}.$$

It follows from Theorem 8.1 with $M = 1$ that $(1 + u(\lambda))^{-1}$ is a completely monotonic function of λ. We have by (6.16),

$$1 + u(\lambda) = 1 - \eta_*(\infty) + \lambda\hat{\eta}_*(\lambda)$$

and
$$\lambda \sum_{j \in \mathbf{I}} \hat{\xi}_j(\lambda) = \frac{\lambda\hat{\eta}_*(\lambda)}{1 + u(\lambda)} = \frac{\lambda\hat{\eta}_*(\lambda)}{\lambda\hat{\eta}_*(\lambda) + 1 - \eta_*(\infty)}.$$

It follows that $\sum_{j\in\mathbf{I}} \xi_j(t) = 1$ or that (p_{ij}), $(i, j) \in \mathbf{I}\times\mathbf{I}$, is stochastic if and only if $\eta_*(\infty) = 1$. In general

$$\lambda\hat{\xi}_\theta(\lambda) = \frac{1 - \eta_*(\infty)}{1 + u(\lambda)}.$$

Hence we have
$$\xi_\theta(0) = \frac{1 - \eta_*(\infty)}{1 - \eta_*(\infty) + \eta_*(0)},$$

$$\xi_\theta(\infty) = 1 - \eta_*(\infty).$$

It is possible to extend the equation (6.7) to ξ_θ as follows. Since

$$\xi_\theta(s + t) - \xi_\theta(s) = \sum_{i\in\mathbf{I}} \xi_i(s)\, p_{i\theta}(t) = \sum_{i\in\mathbf{I}} \xi_i(s) \int_0^t l_i(u)\, \xi_\theta(t - u)\,du,$$

we have
$$\int_0^\infty [\xi_\theta(s + t) - \xi_\theta(s)]\,ds = \int_0^t \sigma(u)\, \xi_\theta(t - u)\,du.$$

Since ξ_θ is absolutely continuous, the left member above is equal to

$$\int_0^\infty \int_s^{s+t} \xi_\theta'(u)\,du\,ds = \int_0^t u\,\xi_\theta'(u)\,du + t\int_t^\infty \xi_\theta'(u)\,du = t\,\xi_\theta(\infty) - \int_0^t \xi_\theta(u)\,du.$$

It follows that
$$t\xi_\theta(\infty) = \int_0^t \xi_\theta(s)\,ds + \int_0^t \sigma(s)\,\xi_\theta(t-s)\,ds.$$

Thus to extend (6.7) to ξ_θ we should set $\eta_\theta(t) \equiv \xi_\theta(\infty)$.

The functions $\eta_j(\cdot)$ in Theorem 6.2 can be decomposed into two parts. Letting $s \downarrow 0$ in (6.5) we have

$$\sum_i \eta_i(0)\,f_{ij}(t) \leqslant \eta_j(t).$$

If we set
$$\bar{\eta}_j(t) = \eta_j(t) - \sum_i \eta_i(0)\,f_{ij}(t),$$

then $\{\bar{\eta}_i(\cdot)\}$ is an entrance solution for Φ satisfying $\bar{\eta}_i(0) = 0$. Consequently, we have by Theorem 6.7,

$$\sum_i \bar{\eta}_i(t)\,q_{ij} = 0.$$

These results check with Reuter [13]. It is to be noted that Reuter's analytical assumption implies our Assumption C, unless I consists of one $\prod$-recurrent class.

The function $m(\lambda)$ in (9.1) is of interest. Note that

$$u(\lambda) = \int_0^\infty \lambda\, e^{-\lambda t}\,dt \int_t^\infty \theta(s)\,ds = \int_0^\infty (1 - e^{-\lambda s})\,\theta(s)\,ds, \tag{9.2}$$

and by (6.7) and Corollary 2 to Theorem 6.3:

$$\int_0^1 s\,\theta(s)\,ds = \int_0^1 ds \int_0^s \sum_i \eta_i(u)\,l_i(s-u)\,du = \int_0^1 \sum_i \eta_i(u)\,L_i(1-u)\,du$$
$$\leqslant \int_0^1 \eta_*(u)\,du \leqslant t + \int_0^t \sigma(s)\,ds < \infty.$$

Hence the last member of (9.2) is the negative Laplace transform of an infinitely divisible distribution on T. Precisely, there is a process $\{Y(v), v \in \mathrm{T}\}$ with stationary independent positive increments such that

$$\mathbf{E}(e^{-\lambda Y(v)}) = \exp\left[v\int_0^\infty (e^{-\lambda s} - 1)\,\theta(s)\,ds\right]. \tag{9.3}$$

It follows that
$$m(\lambda)=\int_0^\infty e^{-v[1+u(\lambda)]}\,dv=\mathbf{E}(e^{-\lambda Y(\mu)}),\tag{9.4}$$

where μ is a random variable with the distribution function e_1 and independent of the process $\{Y(v)\}$.

In the particular case where $\int_0^\infty \theta(s)\,ds=\sigma(0)<\infty$, we set

$$F(t)=\frac{1}{\sigma(0)}\int_0^t \theta(s)\,ds.\tag{9.5}$$

The following theorem is easily proved.

THEOREM 9.1. *Let the random variable μ be as described above and let the random variable ν have the geometric distribution given as follows:*

$$\mathbf{P}\{\nu=n\}=\left(\frac{\sigma(0)}{1+\sigma(0)}\right)^n\frac{1}{1+\sigma(0)},\quad n\in\mathbf{N}.$$

Let $\{y_n, n\in\mathbf{N}\}$ be a sequence of independent random variables having the common distribution function F in (9.5) *and independent of ν. Then $Y(\mu)$ and $\sum_{n=1}^{\nu} y_n$ have the same distribution.*

The matrix generalization of this theorem is implicit in Theorem 8.4; see also the discussion at the end of § 5. For the case where the matrix $(\bar{\sigma}^{aa}(0))$ is infinite on the diagonal and finite elsewhere see Neveu [10], [11]. The extent to which his results generalize Theorem 9.1 is not clear. The representation (9.4) must be intimately related to Paul Lévy's "local time" (see [9]), but again the exact connection is not clear.

10. Appendix

The following theorem, under the additional assumption of (2.4) with equality, was first proved with probabilistic methods by D. G. Austin; a simplified version by the present author is given as Theorem II.5.2 in [1]. A simpler analytic proof was later obtained by D. Ornstein; it is given as Theorem II.1.5 in [1]. The present proof, without the assumption (2.4), is a modification of the latter.

THEOREM 10.1. *Let (p_{ij}), $(i,j)\in\mathbf{I}\times\mathbf{I}$, be a matrix of functions on* $\mathbf{T}$ *satisfying* (2.1), (2.2) *and* (2.3). *Then each $p_{ij}(\cdot)$ is either identically zero or never zero.*

Proof. Suppose $t_0>0$ and $p_{ij}(t_0)=0$. Let N be a positive integer and $t_0=2Ns$. Define

$$C_m = \{k : p_{ik}(ms) = 0\},$$

$$C_m - C_{m+1} = D_{m+1} \quad (m \geqslant 0).$$

Then $C_0 = \mathbf{I} - \{i\}$, $C_m \searrow$, and $j \in C_{2N}$. Let us put

$$u(m, n) = \sum_{k \in C_m} p_{ik}(ns)\, p_{kj}(4Ns - ns),$$

$$v(m, n) = \sum_{k \in D_m} p_{ik}(ns)\, p_{kj}(4Ns - ns).$$

We have $\quad u(m, 0) = 0 \quad (0 \leqslant m); \qquad u(m, 4N) = p_{ij}(4Ns) \quad (0 \leqslant m \leqslant 2N).$

By [3; Theorem 1], each p_{ij} is continuous in $\mathbf{T}$; hence by Dini's theorem, the series

$$\sum_{k \in \mathbf{I}} p_{ik}(t)\, p_{kj}(2t_0 - t) = p_{ij}(2t_0) \tag{10.1}$$

converges uniformly in $t \in [0, 2t_0]$. Since the D_m's are disjoint (possibly void) and

$$\sum_{m=0}^{\infty} v(m, n) \leqslant \sum_{k \in \mathbf{I}} p_{ik}(ns)\, p_{kj}(4Ns - ns),$$

it follows from the uniform convergence of the series in (10.1), that

$$\sum_{m=0}^{\infty} v(m, n) \text{ converges uniformly in } n, \ 0 \leqslant n \leqslant 4N. \tag{10.2}$$

We have by the definitions:

$$u(m, n+1) - v(m+1, n+1) = \sum_{k \in C_{m+1}} \Big(\sum_{l \in \mathbf{I}} p_{il}(ns)\, p_{lk}(s)\Big)\, p_{kj}(4Ns - ns - s). \tag{10.3}$$

If $k \in C_{m+1}$ and $p_{lk}(s) > 0$, then $l \in C_m$; for otherwise $p_{ik}(ms + s) \geqslant p_{il}(ms)\, p_{lk}(s) > 0$ and k would not belong to C_{m+1}. Hence in the double sum in (10.3) we need only sum l over C_m, and consequently

$$u(m, n+1) - v(m+1, n+1) \leqslant \sum_{l \in C_m} p_{il}(ns) \sum_{k \in \mathbf{I}} p_{lk}(s)\, p_{kj}(4Ns - ns - s)$$

$$= \sum_{l \in C_m} p_{il}(ns)\, p_{lj}(4Ns - ns) = u(m, n).$$

Summing over n we obtain

$$p_{ij}(4Ns) = u(m, 4N) \leqslant \sum_{n=0}^{4N-1} v(m+1, n+1).$$

This being true for $0 \leqslant m \leqslant 2N$, we infer that

$$p_{ij}(2t_0) \leqslant \frac{1}{N} \sum_{n=0}^{4N-1} \sum_{m=N}^{2N-1} v(m+1, n+1) \leqslant 4 \max_{1 \leqslant n \leqslant 4N} \sum_{m=N+1}^{2N} v(m, n).$$

As $N \to \infty$, the last member above converges to zero by (10.2), and so $p_{ij}(2t_0) = 0$. Repeating this argument, we see that $p_{ij}(2^n t_0) = 0$ for every positive integer n and consequently $p_{ij}(t) \equiv 0$, since $p_{ij}(t) > 0$ implies $p_{ij}(t') > 0$ for $t' > t$ trivially. The theorem is proved.

COROLLARY. *Let $\{\xi_j(\cdot), j \in \mathrm{I}\}$ be nonnegative functions on T^0 satisfying either*

$$\xi_j(t) = \sum_{i \in \mathrm{I}} \xi_i(s) p_{ij}(t-s) \quad (0 < s < t),$$

or

$$\xi_i(t) = \sum_{j \in \mathrm{I}} p_{ij}(t-s) \xi_j(s) \quad (0 < s < t),$$

for every $t \in \mathrm{T}^0$. Then each $\xi_j(\cdot)$ is either identically zero or never zero in T^0.

Proof. Theorem 10.1 being symmetric in the pair of indices (i, j), we need only prove the first form of the Corollary. If for some $t > 0$ we have $\xi_j(t) > 0$, then for any δ: $0 < \delta < t$, there exist s: $0 < s < \delta$, and $i \in \mathrm{I}$ such that $\xi_i(s) > 0$ and $p_{ij}(t-s) > 0$. Hence by the theorem, $p_{ij}(\delta - s) > 0$ and so $\xi_j(\delta) \geqslant \xi_i(s) p_{ij}(\delta - s) > 0$. Since δ is arbitrary $\xi_j(\cdot) > 0$ in T^0, proving the corollary.

It follows from the Corollary that each function such as k, l, ξ, η, ζ, ψ in the text, which is a member of an exit or entrance solution for a standard transition matrix ($\prod$ or Φ), has the always-or-never-zero property. The result can be generalized at once to a measurable transition matrix (see the last paragraph of p. 122 in [1]).

References

[1]. CHUNG, K. L., *Markov Chains with Stationary Transition Probabilities.* Springer, Berlin Göttingen, Heidelberg, 1960.

[2]. ——, On last exit times. *Illinois J. Math.*, 4 (1960), 629–39.

[3]. ——, *Probabilistic Methods in Markov Chains.* Fourth Berkeley Symposium on Mathematical Statistics and Probability, Vol. II, 35–56, 1961. University of California Press, Berkeley and Los Angeles.

[4]. ——, On the Martin boundary for Markov chains. *Proc. Nat. Acad. Sci.*, 48 (1962), 963–968.

[5]. DOOB, J. L., Discrete potential theory and boundaries. *J. Math. Mech.*, 8 (1959), 433–458.

[6]. FELLER, W., Boundaries induced by positive matrices. *Trans. Amer. Math. Soc.*, 83 (1956), 19–54.

[7]. ——, On boundaries and lateral conditions for the Kolmogorov differential equations. *Ann. of Math.*, 65 (1957), 527–570; "*Notes*", *ibid.*, 68 (1958), 735–736.

[8]. HUNT, G. A., Markoff chains and Martin boundaries. *Illinois J. Math.*, 4 (1960), 313–340.
[9]. LÉVY, P., Systèmes markoviens et stationnaires; cas dénombrable. *Ann. École Norm.* (3), 68 (1951), 327–381.
[10]. NEVEU, J., Une généralisation des processus à accroisements croissants indépendants. *Abh. Math. Mem. Univ. Hamburg*, 25 (1961), 36–61.
[11]. ——, *Lattice methods and submarkovian processes.* Fourth Berkeley Symposium on Mathematical Statistics and Probability, Vol. II, 347–391, 1961. University of California Press, Berkeley and Los Angeles.
[12]. REUTER, G. E. H., Denumerable Markov processes and the associated contraction semiproups on l. *Acta Math.*, 97 (1957), 1–46.
[13]. ——, Denumerable Markov processes (II). *J. London Math. Soc.*, 34 (1959), 81–91.
[14]. ——, Denumerable Markov processes (III). *Ibid.*, 37 (1962), 63–73.

Received March 8, 1963

Z. Wahrscheinlichkeitstheorie 2, 230—254 (1964)

The General Theory of Markov Processes According to Doeblin *

By

KAI LAI CHUNG

§ 0. Introduction

This paper is based on DOEBLIN's paper [1] cited in the Bibliography. Although this work probably represents his crowning achievement in the theory of Markov processes, it is little known and almost never used, even when it is occasionally included in the references as a collector's item. (For what is generally known as DOEBLIN's theory see [2] and [3, Chapter 5].) The present author gave a course on the material of [1] in the spring of 1951 at Columbia University and the lecture notes were mimeographed for limited circulation. The version presented here is an expanded one over these notes, with a number of new results added, but it treats only that part of his theory which may be called the descriptive foundations, stopping short of the principal limit theorem. One reason for doing so is that the presentation of the latter hard theorem still leaves much to be desired, while the part given here seems to have reached a stage where it assumes a quite independent place in the general theory. It is hoped that the appearance in print of this will encourage further research towards various limit theorems in the general context.

It does not seem necessary to detail the differences between this presentation and DOEBLIN's own, since the latter is easily accessible for the sake of comparison. The curious reader may also consult the notes mentioned above which are closer to the original. I shall therefore limit myself to a few remarks. In §§ 1—2 my work has been mainly that of organization and clarification. Proposition 5 is due to BLACKWELL and Proposition 6 to myself, both of which are given new proofs here. Propositions 18 and 19 summarize some basic properties of a specially important type of space; the resemblance of Proposition 18 to the classical theorem of Cantor's on nested sequences of closed sets is notable. §§ 3—4 contain substantial enlargements. In particular in § 3 the arithmetical study in Propositions 34 to 37 (ending in Proposition 45 in § 4) is new. In § 4, Definitions 10 and 11 as well as Propositions 41 and 43 are new, leading to a more stringent definition of our H which resembles DOEBLIN's D, our D being his ϱ. With the present definition the conjecture "$D = H$" was first proved by H. KESTEN (private communcation 1962) and became Proposition 48. In § 5, the proofs of Propositions 50 and 51 are both simpler than DOEBLIN's original ones, the first due to T. E. HARRIS (private communication 1955).

In the remainder of this section we review briefly a constructive definition of Markov processes in the general case considered here. The reader is supposed to

* This research is supported in part by the Office of Scientific Research of the United States Air Force.

have some knowledge of these processes at least in a more limited context. Standard terminology and notation such as in [3] or [4] will be used wherever not specified. The letters i, j, k, l, m, n, ν are positive integers or zero where not specified; the complement of a set is sometimes indicated by the superscript "c".

Let X be an abstract space and $\mathscr{B}$ a Borel field of subsets of X. We are given a function $P(\cdot,\cdot)$ where $x \in X$, $B \in \mathscr{B}$ with the following properties:

(i) for each x, $P(x,\cdot)$ is a probability measure on $\mathscr{B}$;
(ii) for each B, $P(\cdot, B)$ is a $\mathscr{B}$-measurable function of x.

Let furthermore an arbitrary probability measure $P_0(\cdot)$ on $\mathscr{B}$ be given. It is known (see [3; p. 613]) that a probability space $(\Omega, \mathscr{F}, \boldsymbol{P})$ can be constructed to satisfy the following requirements. There exists a sequence of functions $\{\xi_n, n \geqq 0\}$, each of which is from Ω into X and is $(\mathscr{F}, \mathscr{B})$ measurable; that is, $\xi^{-1}(\mathscr{B}) \subset \mathscr{F}$. The measure $\boldsymbol{P}$ is completely determined by $P(\cdot,\cdot)$ and $P_0(\cdot)$ on the Borel subfield $\mathscr{F}_{[0,\infty)}$ generated by $\{\xi_n, n \geqq 0\}$, as follows: for any $B_m \in \mathscr{B}$:

$$\text{(A)} \qquad \boldsymbol{P}\{\xi_m \in B_m, 0 \leqq m \leqq n\} = \int_{B_0} P_0(dx_0) \int_{B_1} P(x_0, dx_1) \cdots \int_{B_m} P(x_{m-1}, dx_m).$$

The sequence $\{\xi_n, n \geqq 0\}$ is a (discrete parameter) Markov process with the stationary transition probability function $P(\cdot,\cdot)$ and the initial distribution P_0.

In the particular case $P_0(\cdot) = \delta(x,\cdot)$ where for every $B \in \mathscr{B}$:

$$\delta(x, B) = \begin{cases} 1 & \text{if } x \in B, \\ 0 & \text{if } x \notin B; \end{cases}$$

the corresponding $\boldsymbol{P}$ restricted to $\mathscr{F}_{[0,\infty)}$ will be denoted by $\boldsymbol{P}_x$, and the corresponding Markov process is said to start from x. The function $\boldsymbol{P}_x\{\Lambda\}$ where $x \in X$, $\Lambda \in \mathscr{F}_{[0,\infty)}$ is useful since it is a well-defined and convenient version of the conditional probability $\boldsymbol{P}\{\Lambda | \xi_0 = x\}$.

More generally, let $\mathscr{F}_{[n,\infty)}$ be the Borel subfield of $\mathscr{F}$ generated by $\{\xi_m, m \geqq n\}$, then for each $n \geqq 0$ and each $\Lambda \in \mathscr{F}_{[n,\infty)}$, we have for every ω except a set of $\boldsymbol{P}$-measure zero:

$$\text{(B)} \qquad \boldsymbol{P}\{\Lambda \mid \xi_0(\omega), \ldots, \xi_n(\omega)\} = \boldsymbol{P}_{\xi_n(\omega)}\{\Lambda\},$$

where $\boldsymbol{P}$ is given by (A) with an arbitrary P_0. The Markov property of the process is embodied in the equation (B).

Several cases of $\boldsymbol{P}_x(\Lambda)$ for important sets Λ will now be given with special symbols assigned to them. These will be employed throughout the paper and simple intuitive relations connecting them based on the above interpretations of conditional probabilities will be passed muster.

We write $B^c = X - B$ below:

$P^{(n)}(x,B) = \boldsymbol{P}_x\{\xi_n \in B\}$ for $n \geqq 0$ is obtained by putting $B_0 = \{x\}$, $B_1 = \cdots = B_{n-1} = X$, and $B_n = B$ in formula (A);

$K^{(n)}(x,B) = \boldsymbol{P}_x\{\xi_m \in B^c, 1 \leqq m \leqq n-1; \xi_n \in B\}$ for $n \geqq 1$ is obtained by putting $B_0 = \{x\}$, $B_1 = \cdots = B_{n-1} = B^c$, and $B_n = B$ in formula (A);

$$L(x,B) = \sum_{n=1}^{\infty} K^{(n)}(x,B) = \boldsymbol{P}_x\{\bigcup_{n=1}^{\infty} [\xi_n \in B]\};$$

$$Q(x, B) = 1 - \sum_{n=0}^{\infty} \int_B P^{(n)}(x, dy)[1 - L(y, B)]$$
$$= \lim_{n\to\infty} \int_X P^{(n)}(x, dy)\, L(y, B)$$
$$= \boldsymbol{P}_x\{\bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} [\xi_n \in B]\}.$$

Note that $P^{(0)}(x,B) = \delta(x,B)$; and $P^{(1)}(x,B) = P(x,B)$.

§ 1. Closedness and essentialness

For an arbitrary set E in $\mathscr{B}$, we define four sets:

$$E^0 = \{x : L(x,E) = 0\},$$
$$E^1 = \{x : L(x,E) = 1\},$$
$$E^f = \{x : Q(x,E) = 0\},$$
$$E^\infty = \{x : Q(x,E) = 1\}.$$

For each E in $\mathscr{B}$, the functions $L(\cdot,E)$ and $Q(\cdot,E)$ are $\mathscr{B}$-measurable; hence each of the four sets above is in $\mathscr{B}$. The complement of E^*, where $*$ stands for any of the symbols 0, 1, f or ∞, will be denoted by E^{*c} rather than $(E^*)^c$.

Definition 1. A nonempty set E in $\mathscr{B}$ such that $P(x,E) = 1$ for every $x \in E$ is called *stochastically closed* (cl.).

Proposition 1. *If $x \in E^0$, then $P(x, E^0E^c) = 1$. The sets E^0 and E^0E^c are either both empty or both cl.*

Proof. We have

$$0 = L(x,E) = \int_E P(x,dy) + \int_{E^0E^c} L(y,E)\, P(x,dy) + \int_{E^{0c}E^c} L(y,E)\, P(x,dy).$$

Since the integrands in the first and third integrals above are positive, we have

$$P(x, E \cup E^{0c}E^c) = 0$$

or

$$P(x, E^0E^c) = 1.$$

It follows that E^0 as well as E^0E^c is cl. unless empty, and that if E^0 is nonempty then so is E^0E^c.

Proposition 2. *If $x \in E^1$, then $P(x, E^1 \cup E) = 1$. If E is cl., then so is E^1.*

Proof. We have

$$1 = L(x,E) = P(x,E^1) + P(x,E^{1c}E) + \int_{E^{1c}E^c} L(y,E)\, P(x,dy).$$

Since the integrand in the last integral is less than 1, we have

$$P(x, E^{1c}E^c) = 0$$

or

$$P(x, E^1 \cup E) = 1.$$

If E is cl., then $E \subset E^1$; hence the first assertion implies the second.

Proposition 3. *E^f is cl. or empty.*

Proof. Suppose E^f is nonempty and let $x \in E^f$; then

$$0 = Q(x,E) = \int_X Q(y,E)\, P(x,dy) \geqq \int_{E^{fc}} Q(y,E)\, P(x,dy) \geqq 0$$

Since the integrand in the last integral is positive, we have $P(x,E^{fc}) = 0$ or $P(x,E^f) = 1$.

Proposition 4. *E^∞ is cl. or empty.*

Proof. Suppose E^∞ is nonempty and let $x \in E^\infty$; then

$$1 = Q(x,E) = \int_{E^\infty} Q(y,E)\, P(x,dy) + \int_{E^{\infty c}} Q(y,E)\, P(x,dy)\,.$$

Since the integrand in the last integral is less than 1, we have $P(x,E^{\infty c}) = 0$ or $P(x,E^\infty) = 1$.

Proposition 5. *If $E = \bigcup_n E_n$, then $E^0 = \bigcap_n E_n^0$; where $\{E_n\}$ is an arbitrary sequence of sets in $\mathscr{B}$.*

Proof. Clearly $E^0 \subset E_n^0$ so that $E^0 \subset \bigcap_n E_n^0$. On the other hand, if $x \in \bigcap_n E_n^0$, then for every n we have $L(x,E_n) = 0$; consequently

$$L(x,E) = L(x, \bigcup_n E_n) \leqq \sum_n L(x,E_n) = 0\,,$$

and so $x \in E^0$. Thus $\bigcap_n E_n^0 \subset E^0$.

Definition 2. A set E in $\mathscr{B}$ such that $Q(x,E) = 0$ for every $x \in X$ is called *inessential (iness.)*; otherwise it is called *essential (ess.)*. An essential set which is the union of denumerably many inessential sets is called *improperly essential (imp. ess.)*; otherwise it is called *absolutely essential (abs. ess.)*.

The next two propositions are basic for the sequel. Proposition 6 was given by Blackwell [*5*], and Proposition 7 in the lecture notes mentioned in the Introduction and essentially reproduced in [*4*; p. 19]. Both were proved by simple, direct arguments. For the sake of completeness but variation we give alternative proofs below based on the convergence of martingales.

For any E in $\mathscr{B}$, let

$$\Lambda(E) = \limsup_n \{\xi_n \in E\}\,,$$

and for any E and F in $\mathscr{B}$, let

$$Q(x,E,F) = \boldsymbol{P}_x\{\Lambda(E) \cap \Lambda(F)\}\,.$$

Proposition 6. *If*

$$\sup_{x \in E} Q(x,F) < 1\,,$$

then for every $x \in X$ we have

$$Q(x,E,F) = 0$$

Proof. Fix an x as the initial point of the process $\{\xi_n, n \geqq 0\}$. Since $\Lambda(E)$ and $\Lambda(F)$ are invariant sets we have with probability one:

$$\boldsymbol{P}_x\{\Lambda(E) \cap \Lambda(F) \mid \xi_0, \ldots, \xi_n\} = \boldsymbol{P}_x\{\Lambda(E) \cap \Lambda(F) \mid \xi_n\} = Q(\xi_n, E, F)\,.$$

This shows that $\{Q(\xi_n, E, F), n \geqq 0\}$ is a martingale and PAUL LÉVY's zero-or-one law asserts that for almost every ω:

$$\lim_{n\to\infty} Q(\xi_n(\omega), E, F) = \Im_{\Lambda(E)\cap\Lambda(F)} \tag{1}$$

where $\Im_\Lambda$ denotes the indicator function of Λ. Now if $\omega \in \Lambda_E$, then $\xi_n(\omega) \in E$ for infinitely many values of n, and consequently for these values of n we have

$$Q(\xi_n(\omega), E, F) \leqq Q(\xi_n(\omega), F) \leqq \sup_{x\in E} Q(x, F) < 1\,. \tag{2}$$

It follows from (1) and (2) that $P_x(\Lambda(E) \cap \Lambda(F)) = 0$.

Proposition 7. *If*

$$\inf_{x\in E} L(x, F) > 0\,,$$

then for every $x \in X$ we have

$$Q(x, E) = Q(x, E, F)\,.$$

Proof. Let

$$M_k = \bigcup_{n=k}^{\infty}\{\xi_n \in F\}$$

so that in the previous notation we have

$$\bigcap_{k=1}^{\infty} M_k = \Lambda(F)\,.$$

We have if $n \geqq k$:

$$P_x\{\Lambda_F|\xi_0, \ldots, \xi_n\} \leqq P_x\{M_{n+1}|\xi_0, \ldots, \xi_n\} \leqq P_x\{M_k|\xi_0, \ldots, \xi_n\}$$

with probability one. Letting $n \to \infty$, then $k \to \infty$, we obtain

$$\Im_{\Lambda(F)} \leqq \lim_{n\to\infty} P_x\{M_{n+1}|\xi_0, \ldots, \xi_n\} \leqq \lim_{k\to\infty} \Im_{M_k} = \Im_{\Lambda(F)}\,.$$

Since

$$L(\xi_n, F) = P_x\{M_{n+1}|\xi_n\} = P_x\{M_{n+1}|\xi_0, \ldots, \xi_n\}$$

with probability one, we conclude that

$$\varlimsup_{n\to\infty} L(\xi_n, F) = \Im_{\Lambda(F)}\,. \tag{3}$$

If $\omega \in \Lambda(E)$, then $\xi_n(\omega) \in E$ for infinitely many values of n and consequently for these values of n we have

$$L(\xi_n(\omega), F) \geqq \inf_{x\in E} L(x, F) > 0\,. \tag{4}$$

It follows from (3) and (4) that $P_x(\Lambda(E)) = P_x(\Lambda(E) \cap \Lambda(F))$.

Proposition 8. *If E is ess., and* $\inf_{x\in E} L(x, F) > 0$, *then F is ess.*

Proof. Since E is ess. there exists an x for which $Q(x, E) > 0$. By Proposition 7,

$$Q(x, F) \geqq Q(x, E, F) = Q(x, E) > 0\,.$$

Hence F is ess.

Proposition 9. *If E is abs. ess., and $\inf_{x \in E} L(x, F) > 0$, then F is abs. ess.*

Proof. It is sufficient to prove that for any sequence of sets F_k in $\mathscr{B}$ such that $F = \bigcup_{k=1}^{\infty} F_k$, there exists an n_0 such that $\bigcup_{k=1}^{n_0} F_k$ is ess. We note the simple relation:

$$L(x, F) = \lim_{n \to \infty} L(x, \bigcup_{k=1}^{n} F_k).$$

Let $x \in E$, then $L(x, F) = \alpha > 0$; hence there exists a finite $m_0(x)$ such that

$$L(x, \bigcup_{k=1}^{m_0(x)} F_k) > \frac{\alpha}{2} > 0.$$

Let $E_n = \{x \in E : m_0(x) = n\}$, then $E = \bigcup_{n=1}^{\infty} E_n$. Since E is abs. ess. there exists an n_0 such that E_{n_0} is ess. By the definition of E_{n_0} we have

$$\inf_{x \in E_{n_0}} L(x, \bigcup_{k=1}^{n_0} F_k) > \frac{\alpha}{2}.$$

Hence by Proposition 8, $\bigcup_{k=1}^{n_0} F_k$ is ess.

Proposition 10. *For any E in $\mathscr{B}$, if there exists an F in $\mathscr{B}$ such that*

$$\sup_{x \in E} Q(x, F) < 1, \qquad \inf_{x \in E} L(x, F) > 0,$$

then E is iness.

Proof. For every x we have by Propositions 6 and 7:

$$Q(x, E) = Q(x, E, F) = 0$$

Hence E is iness. by definition.

Proposition 11. *If $X - E^0$ is abs. ess., then E is abs. ess.*

Proof. Let

$$E_n = \left\{x : L(x, E) \geqq \frac{1}{n}\right\};$$

then we have

$$X = E^0 \cup \bigcup_{n=1}^{\infty} E_n.$$

If $X - E^0$ is abs. ess., then E_n is abs. ess. for some n, and so E is abs. ess. by Proposition 9.

Proposition 11.1. *If X is abs. ess. and $E^0 = 0$, then E is abs. ess.*

Proposition 12. *For any E in $\mathscr{B}$, $X - (E^0 \cup E^\infty)$ is not abs. ess.*

Proof. Let

$$E_n = \left\{x : Q(x, E) \leqq 1 - \frac{1}{n}, L(x, E) \geqq \frac{1}{n}\right\};$$

then

$$X = E^0 \cup E^\infty \cup \bigcup_{n=1}^{\infty} E_n.$$

Each E_n is iness. by Proposition 10, hence their union $X - (E^0 \cup E_\infty)$ is not abs. ess.

Proposition 12.1. *If E is abs. ess., then for any F in $\mathcal{B}$ the set $E(F^0 \cup F^\infty)$ is abs. ess., hence nonempty.*

Proposition 13. *If E is abs. ess., then EE^∞ is abs. ess., in particular $E^\infty \neq 0$.*

Proof. Applying Proposition 12.1 with $F = E$ we see that $EE^0 \cup EE^\infty$ is abs. ess. But EE^0 is clearly iness., hence EE^∞ is abs. ess.

Proposition 14. *If C is cl., then $X - (C \cup C^0)$ does not contain any cl. set and is not abs. ess.*

Proof. Any cl. set contained in $X - C$ must be contained in C^0, hence $E = X - (C \cup C^0)$ does not contain any cl. set. Since $C \cup C^0$ is cl., any point in E^1 must belong to E; in particular $E^\infty \subset E^1 \subset E$. But E^∞ is cl. if not empty, hence $E^\infty = 0$ by the first assertion. It follows from Proposition 13 that E is not abs. ess.

Proposition 14.1. *If C and D are cl. sets such that $D \subset C$ and $C - D$ does not contain any cl. set, then $C - D$ is not abs. ess. In particular, $C - C(D \cup D^0)$ is not abs. ess.*

§ 2. Decomposability

Definition 3. A cl. set which does not contain two disjoint cl. sets is called *indecomposable (indecomp.)*; otherwise it is called *decomposable (decomp.)*. An indecomposable set which is not properly contained in any indecomposable set is called *maximal indecomposable (max. indecomp.)*

Proposition 15. *If E is indecomp., then $(E^0)^0$ is max. indecomp.*

Proof. Suppose $(E^0)^0$ is decomp.; let C and D be two disjoint cl. sets contained in it. For any $x \in C$ we have $x \notin E^0$ since $E^0 (E^0)^0 = 0$; hence $L(x, E) > 0$. Since C is cl. this implies that $CE \neq 0$. Similarly $DE \neq 0$. Thus CE and DE are disjoint cl. sets contained in E and E is decomp. We have thus proved that if E is indecomp., then so is $(E^0)^0$. Now suppose that F is cl. and contains $(E^0)^0$ properly. Let $x \in F - (E^0)^0$, then $L(x, E^0) > 0$. Thus E^0 is nonempty and hence cl. by Proposition 1, and FE^0 is also nonempty and hence cl. The set F contains the disjoint cl. sets E and FE^0 and so is decomp. We have therefore proved that any cl. set properly containing $(E^0)^0$ is decomp. Hence $(E^0)^0$ is max. indecomp.

Proposition 16. *Two max. indecomp. sets are either identical or disjoint.*

Proof. Let E and F be two distinct max. indecomp. sets. Then $E \cup F$ is cl. and contains either of them properly. Hence it is decomp. and contains two disjoint cl. sets C and D. Since E is indecomp. at least one of EC and ED is empty. Suppose EC is empty; then $F \supset C$ and since F also contains EF which is either cl. or empty we must have $EF = 0$ since F is indecomp.

Proposition 17. *If X is indecomp. and E is abs. ess., then $E^0 = 0$.*

Proof. By Proposition 13 we have $E^\infty \neq 0$, hence E^∞ is cl. by Proposition 4. By Proposition 1, E^0 is cl. if not empty. Since $E^0 E^\infty = 0$ and X is indecomp., we must have $E^0 = 0$.

Proposition 18. *If X is abs. ess. and indecomp., then every sequence of cl. sets has a cl. intersection whose complement is not abs. ess.*

Proof. Let $\{C_k\}$ be a finite or infinite sequence of cl. sets. Then $D_n = \bigcap_{k=1}^{n} C_k$ is not empty since X is indecomp. We have

$$X = \bigcup_n D_n^c \cup (\bigcap_n D_n).$$

Each D_n is cl. and D_n^c does not contain any cl. set by the indecomposability of X. Hence D_n^c is not abs. ess. by Proposition 14.1. with $C = X$; and so $\bigcup_n D_n^c$ is not abs. ess. Since X is abs. ess. it follows that $\bigcap_n D_n$ is abs. ess., hence it is nonempty, hence it is cl.

Proposition 18.1. *In an indecomp. space the complement of any cl. set is not abs. ess.*

Proposition 19. *In an abs. ess. and indecomp. space X, an abs. ess. set E is characterized by any one of the following three properties:*

$$E^0 = 0, \quad E^\infty \neq 0, \quad E^f = 0.$$

Proof. The first characterization follows from Propositions 11.1 and 17. Next, each of the three sets E^0, E^∞ and E^f is either cl. or empty, by Propositions 1, 4 and 3. Now at least one of the two sets E^0 and E^∞ is nonempty by Proposition 12.1. Hence exactly one of them is nonempty since $E^0 E^\infty = 0$ and X is indecomp. Thus $E^\infty \neq 0$ is equivalent to $E^0 = 0$ and we have proved the second characterization. Finally, since $E^0 \subset E^f$, $E^f = 0$ implies $E^0 = 0$; on the other hand since $E^f E^\infty = 0$, $E^f \neq 0$ implies $E^\infty = 0$ because of indecomposability. Hence the third characterization is a consequence of the first two.

Remark: Let "$E \in \mathscr{A}$" stand for the proposition "E is abs. ess.", "$\Rightarrow$" for "implies" and "$\not\Rightarrow$" for "does not imply". The following table shows the various relations under different hypotheses regarding the space X; where "$\not\Rightarrow$" stands the required example is trivial from the theory of Markov chains.

Table

Arbitrary X	Abs. ess. X	Indecomp. X	Abs. ess. and indecomp. X
$E \in \mathscr{A} \Rightarrow E^\infty \neq 0$	$E \in \mathscr{A} \not\Rightarrow E^0 = 0$ $E \in \mathscr{A} \not\Rightarrow E^f = 0$ $E^\infty \neq 0 \not\Rightarrow E \in \mathscr{A}$ $E^0 = 0 \Rightarrow E \in \mathscr{A}$ $E^f = 0 \Rightarrow E \in \mathscr{A}$	$E \in \mathscr{A} \Rightarrow E^0 = 0$ $E \in \mathscr{A} \Rightarrow E^f = 0$ $E^\infty \neq 0 \not\Rightarrow E \in \mathscr{A}$ $E^0 = 0 \not\Rightarrow E \in \mathscr{A}$ $E^f = 0 \not\Rightarrow E \in \mathscr{A}$	$E^\infty \neq 0 \Rightarrow E \in \mathscr{A}$

Proposition 20. *If X is indecomp. and E is abs. ess., then the series*

$$\sum_{n=0}^{\infty} P^{(n)}(x, E) \tag{5}$$

diverges for every $x \in X$. If X is abs. ess. and the series in (5) *has a positive sum for every $x \in X$, then E is abs. ess.*

Proof. Suppose that the series in (5) converges for some x, then by the Borel-Cantelli lemma: $Q(x, E) = 0$ so that $E^f \neq 0$. If X is indecomp. a glance at the preceding table shows that E is not abs. ess. Next suppose that the series in (5)

has a positive sum for every x, then $E^0 = 0$. If X is abs. ess., a glance at the preceding table shows that E is abs. ess.

Remark. The converse to the first assertion in Proposition 20 is false. More precisely, it is possible in an indecomp. X that the series in (5) diverges for every x but E is iness. Consider the following example from Markov chains. The space X consists of $\{y_n, n \geqq 1\}$ and $\{x_{nk}, 1 \leqq k \leqq n, n \geqq 1\}$.

$$P(y_1, y_2) = \frac{1}{2};$$

$$P(y_n, y_{n+1}) = 1 - \frac{1}{n^2}; \qquad P(y_n, y_1) = \frac{1}{n^2}, \qquad n \geqq 2;$$

$$P(y_1, x_{n1}) = p_n = \frac{3}{\pi^2 n^2}, \qquad n \geqq 1;$$

$$P(x_{nk}, x_{n,k+1}) = 1, \qquad 1 \leqq k \leqq n-1;$$

$$P(x_{n,n}, y_1) = 1.$$

It is clear that X forms one nonrecurrent class. Let

$$E = \{x_{nk}, 1 \leqq k \leqq n, n \geqq 1\}.$$

We have

$$P^{(n)}(y_1, E) \geqq \sum_{k=n}^{\infty} p_k = \frac{3}{\pi^2} \sum_{k=n}^{\infty} \frac{1}{k^2},$$

so that the series in (5) diverges for $x = y_1$. Since $L(x, y_1) > 0$ for every x it follows easily that it diverges for every x. To see that E is iness., we verify that

$$\inf_{x \in E} L(x, x_{11}) = L(y_1, x_{11}) = p_1 > 0,$$

$$\sup_{x \in E} Q(x, x_{11}) = Q(y_1, x_{11}) \leqq L(y_1, x_{11}) \leqq 1 - \frac{1}{2} \prod_{n=2}^{\infty} \left(1 - \frac{1}{n^2}\right) < 1.$$

Hence E is iness. by Proposition 10.

Definition 4. A set E in $\mathscr{B}$ such that $Q(x, E^c) < 1$ for some $x \in X$ is called *perpetuable (perp.)*. Equivalently, E in $\mathscr{B}$ is perpetuable if there exists an x in E such that $L(x, E^c) < 1$ (see Proposition 24 below). In the literature a perp. set has been called a "sojourn set"; cf. [*4*; p. 110].

Proposition 21. *If E is perp. then it is ess.*

Proof. We have for every E in $\mathscr{B}$:

$$Q(x, E) + Q(x, E^c) \geqq 1, \tag{6}$$

hence $Q(x, E^c) < 1$ implies $Q(x, E) > 0$.

Remark. A set E for which there is equality in (6) for every $x \in X$ has been called "almost closed"; cf. [*4*; p. 108].

Proposition 22. *If C is cl. and C^c is ess., then C^c is perp.*

Proof. Since C^c is ess. there exists an x for which $Q(x, C^c) > 0$. Since C is cl. this implies that

$$Q(x, C) \leqq L(x, C) < 1.$$

Hence C^c is perp.

Proposition 23. *Any imp. ess. set is contained in an imp. ess. and perp. set.*

Proof. If X is imp. ess. the proposition is trivial. Now suppose that X is abs. ess. and E is imp. ess. Then $E^0 \neq 0$ by Proposition 11.1, and $E^0 E^c$ is cl. by Proposition 1. Since E is not abs. ess., E^{0c} is not abs. ess. by Proposition 11. Hence

$$E^{0c} \cup E = (E^0 E^c)^c$$

is not abs. ess. It contains E and is perp. by Proposition 22 since it is ess. and its complement is closed.

Proposition 23.1. *If $E \subset C$ where E is imp. ess. and C is cl., then there exists an imp. ess. and perp. F such that $E \subset F \subset C$ and $C - F$ is cl.*

Proposition 24. *If E is perp. then*

$$\inf_{x \in E} L(x, E^c) = 0 .$$

Proof. We have for any x in X and E in $\mathscr{B}$, as a completion of (6):

$$(6^*) \qquad 1 = Q(x, E \cup E^c) = Q(x, E) + Q(x, E^c) - Q(x, E, E^c) .$$

If $\inf_{x \in E} L(x, E^c) > 0$ then by Proposition 7 we have

$$Q(x, E) = Q(x, E, E^c)$$

so that the equation (6*) implies $Q(x, E^c) = 1$ for every $x \in X$. Thus E is not perp.

§ 3. Cycles

The properties of a set $\mathscr{B}$ such as "closed" and "essential" were defined with reference to the basic transition probability function $P(\cdot,\cdot)$. If the latter is replaced by its k^{th} iterate $P^{(k)}(\cdot,\cdot)$ then the corresponding property will be prefixed by "$P^{(k)}$ -". Thus the previously defined concepts are the $P^{(1)}$-versions, with the prefix "$P^{(1)}$ -" omitted from the terminology. The results we have proved so far have their $P^{(k)}$-versions which need no new proofs. In terms of the process, we shall be considering $\{\xi_{nk+r}, n \geqq 0\}$ for a fixed k and some r in lieu of $\{\xi_n, n \geqq 0\}$.

Proposition 25. *A set is $P^{(k)}$-iness., $P^{(k)}$-imp. ess., or $P^{(k)}$-abs. ess. according as it is iness., imp. ess., or abs. ess.*

Proof. If a set is iness., it is clearly $P^{(k)}$-iness. If E is ess. there exists an $x \in X$ such that $Q(x, E) > 0$. Then for each k there exists an r, $1 \leq r \leq k$, such that

$$P\{\xi_{nk+r} \in E \text{ for infinitely many values of } n \mid \xi_0 = x\} > 0.$$

Hence there exists a $y \in E$ and an integer n_0 such that

$$P\{\xi_{nk+r} \in E \text{ for infinitely many values of } n \mid \xi_{n_0 k+r} = y\} > 0.$$

This shows that E is $P^{(k)}$-ess. The other assertions follow easily.

Definition 5. For an arbitrary set E in $\mathscr{B}$ we set

$$\mathfrak{A}(E) = \{x : P(x, E) = 1\} .$$

Let $\mathfrak{A}^0(E) = E$, $\mathfrak{A}^1(E) = \mathfrak{A}(E)$ and define $\mathfrak{A}^j(E)$ for each $j \geqq 1$ by

$$\mathfrak{A}^j(E) = \mathfrak{A}(\mathfrak{A}^{j-1}(E)).$$

$\mathfrak{A}^j(E)$ is called the j^{th} *antecedent* of E.

We have $\mathfrak{A}(E) \in \mathscr{B}$ for any $E \in \mathscr{B}$, since $P(\cdot, E)$ is $\mathscr{B}$-measurable for each $E \in \mathscr{B}$.

Proposition 26. *If $P^{(k)}(x, E) = 1$ then we have*

$$P^{(k-j)}(x, \mathfrak{A}^j(E)) = 1, \qquad 1 \leqq j \leqq k.$$

Proof. We have

$$1 = P^{(k)}(x, E) = \Big[\int_{\mathfrak{A}(E)} + \int_{\mathfrak{A}(E)^c}\Big] P(y, E)\, P^{(k-1)}(x, dy),$$

where in the second integral the integrand is less than one. Hence the assertion follows for $j = 1$, and the general case then follows from this by induction on j.

Proposition 27. *We have for each $j \geqq 0$,*

$$\mathfrak{A}^j(E) = \{x : P^{(j)}(x, E) = 1\}.$$

Proof. The assertion is true for $j = 1$ by definition. Assume for the sake of induction that it is true for a certain j, then if $x \in \mathfrak{A}^{j+1}(E) = \mathfrak{A}(\mathfrak{A}^j(E))$ we have $P(x, \mathfrak{A}^j(E)) = 1$ and consequently

$$P^{(j+1)}(x, E) = \int_{\mathfrak{A}^j(E)} P^{(j)}(y, E)\, P(x, dy) = \int_{\mathfrak{A}^j(E)} 1\; P(x, dy) = 1$$

by the induction hypothesis. Hence $\mathfrak{A}^{j+1}(E) \subset \{x : P^{(j+1)}(x, E) = 1\}$ by induction. Conversely, if $P^{(j+1)}(x, E) = 1$ then by Proposition 26 we have $P^{(1)}(x, \mathfrak{A}^j(E)) = 1$, and so by definition $x \in \mathfrak{A}(\mathfrak{A}^j(E)) = \mathfrak{A}^{j+1}(E)$.

Definition 6. A sequence of k sets $\{E_j, 1 \leqq j \leqq k\}$ in $\mathscr{B}$ is said to form a *k-cycle* if

$$E_j \subset \mathfrak{A}(E_{j+1}), \qquad 1 \leqq j \leqq k - 1,$$

and

$$E_k \subset \mathfrak{A}(E_1).$$

The union $\bigcup_{j=1}^{k} E_j$ will also be called the *cycle* when no confusion is likely and each E_j, $1 \leqq j \leqq k$, a *member* of the cycle. The cycle is called *clean* if the E_j's are disjoint. Note that in general the members of a cycle need not be distinct.

Proposition 28. *Each member of a k-cycle is $P^{(k)}$-cl. and the cycle itself is $P^{(1)}$-cl. If E is $P^{(k)}$-cl., then the sequence $\mathfrak{A}^{k-j}(E)$, $1 \leqq j \leqq k$, forms a cycle.*

Proof. It follows from Proposition 27 that $E \subset \mathfrak{A}^k(E)$ if and only if E is $P^{(k)}$-cl. Now if $E \subset F$ then $\mathfrak{A}(E) \subset \mathfrak{A}(F)$. Hence by the definition of a cycle we have

$$E_j \subset \mathfrak{A}(E_{j+1}) \subset \cdots \subset \mathfrak{A}^{k-j}(E_k) \subset \mathfrak{A}^{k-j+1}(E_1) \subset \cdots \subset \mathfrak{A}^k(E_j).$$

Thus each E_j is $P^{(k)}$-cl. Furthermore we have

$$\bigcup_{j=1}^{k} E_j \subset \bigcup_{j=0}^{k-1} \mathfrak{A}(E_{j+1}) = \mathfrak{A}\Big(\bigcup_{j=1}^{k} E_j\Big);$$

hence the cycle is $P^{(1)}$-cl.

If E is $P^{(k)}$-cl., then

$$\mathfrak{A}^0(E) = E \subset \mathfrak{A}^k(E) = \mathfrak{A}(\mathfrak{A}^{k-1}(E)),$$

and

$$\mathfrak{A}^{k-j}(E) = \mathfrak{A}(\mathfrak{A}^{k-j-1}(E)), \qquad 1 \leqq j \leqq k;$$

and consequently $\{\mathfrak{A}^{k-j}(E), 1 \leqq j \leqq k\}$ forms a cycle.

Definition 7. The cycle in the second part of Proposition 28 is said to be *generated by* E.

Proposition 29. *If* E *is* $P^{(k)}$-cl. *or* $P^{(k)}$-*indecomp. or* $P^{(k)}$-*max. indecomp., then so is* $\mathfrak{A}^j(E)$ *for each* $j \geqq 0$.

Proof. It is sufficient to prove the assertion for $j = 1$, since the general case then follows by iteration. Let $F = \mathfrak{A}(E)$. If E is $P^{(k)}$-cl., then F is $P^{(k)}$-cl. by the second part of Proposition 28. Next, suppose that E is $P^{(k)}$-cl. and F is $P^{(k)}$-decomp; we are going to show that E is $P^{(k)}$-decomp. There exist disjoint $P^{(k)}$-cl. subsets F_1 and F_2 of F. Define

$$E_n = \mathfrak{A}^{k-1}(F_n), \qquad n = 1, 2.$$

Then E_1 and E_2 are disjoint $P^{(k)}$-cl. sets. If $x \in F_n$, we have

$$1 = P^{(k)}(x, F_n) = \int_E P^{(k-1)}(y, F_n)\, P(x, dy)$$

since $P(x, E) = 1$ by the definition of F. Hence there exists a $y \in E$ with $P^{(k-1)}(y, F_n) = 1$ and consequently $y \in E_n$ by definition. Thus $E \cap E_n \neq 0$ for $n = 1, 2$. Each $E \cap E_n$ is $P^{(k)}$-cl. and so E is $P^{(k)}$-decomp. as was to be shown.

Finally, suppose that E is $P^{(k)}$-max. indecomp. Then F is $P^{(k)}$-indecomp. as just proved. Let $\widetilde{F}$ be $P^{(k)}$-cl. and contain F properly. Define $\widetilde{E} = \mathfrak{A}^{k-1}(\widetilde{F})$. Then $\widetilde{E} \supset E$. If $x \in \widetilde{F}$ then $P(x, \widetilde{E}) = 1$ by Proposition 26. If $x \notin F$ then $P(x, E) < 1$ by the definition of F. Since $\widetilde{F} - F$ is nonempty we see by choosing an x in this difference that $\widetilde{E}$ contains E properly. Hence $\widetilde{E}$ is $P^{(k)}$-decomp. and so must be $\widetilde{F}$ by what has been proved. Therefore F is $P^{(k)}$-max. indecomp.

Notation. If k_1 and k_2 are two positive integers, we write $k_1 | k_2$ if k_1 is a divisor of k_2.

Proposition 30. *Let* $d | k$. *A* $P^{(d)}$-*cl. set is* $P^{(k)}$-*cl. A* $P^{(k)}$-*cl. and* $P^{(k)}$-*indecomp. set is* $P^{(d)}$-*indecomp. A* $P^{(d)}$-*max. indecomp. and* $P^{(k)}$-*indecomp. set is* $P^{(k)}$-*max. indecomp. A* $P^{(d)}$-*cl. and* $P^{(k)}$-*max. indecomp. set is* $P^{(d)}$-*max. indecomp.*

Proof. Without loss of generality we may suppose $d = 1$, since we may consider $P^{(d)}(\cdot,\cdot)$ in lieu of $P(\cdot,\cdot)$ as the basic transition probability function. The first two assertions are trivial.

Let E be $P^{(1)}$-max. decomp. and $P^{(k)}$-indecomp. and let F be a $P^{(k)}$-cl. set which contains E properly. We are going to show that F is $P^{(k)}$-decomp. Let G be the k-cycle generated by F. Then G is $P^{(1)}$-cl. and contain E properly. Hence G contains two disjoint $P^{(1)}$-cl. sets A and B. If $x \in A$ then by the defining property of a cycle we have $P^{(j)}(x, F) = 1$ for some j, $1 \leqq j \leqq k$. Since A is $P^{(1)}$-cl. this implies $A \cap F \neq 0$. By the same token $B \cap F \neq 0$. The two sets $A \cap F$ and $B \cap F$ are disjoint and $P^{(k)}$-cl. Hence F is $P^{(k)}$-decomp. as was to be shown.

To prove the last assertion in Proposition 30, let E be $P^{(1)}$-cl. and $P^{(k)}$-max. indecomp. Then E is $P^{(1)}$-indecomp. by the second assertion in Proposition 30. Let F be $P^{(1)}$-cl. and contain E properly; we are going to show that F is $P^{(1)}$-decomp. Since F is $P^{(k)}$-cl. it must be $P^{(k)}$-decomp. Let A and B be disjoint, $P^{(k)}$-cl. sets contained in F. Since E is $P^{(k)}$-indecomp. at least one of $A \cap E$ and $B \cap E$ is empty. Suppose $A \cap E = 0$ and let C be the k-cycle generated by A. Then $C \cap E = 0$ by the property of a cycle. Hence C and E are disjoint, $P^{(1)}$-cl. sets contained in F and F is $P^{(1)}$-decomp. as was to be shown.

In Propositions 31 *to* 37 *the state space X is assumed to be $P^{(1)}$-indecomp.*

Proposition 31. *There are at most k disjoint $P^{(k)}$-cl. sets.*

Proof. Let B_m, $1 \leq m \leq n$, be disjoint, $P^{(k)}$-cl. sets. By Proposition 28, each of them generates a k-cycle C_m which is $P^{(1)}$-cl. Since X is $P^{(1)}$-indecomp. $C = \bigcap_{m=1}^{n} C_m$ is nonempty. Let $x \in C$, then by the property of a k-cycle for each m, $1 \leq m \leq n$, there exists an integer j_m, $1 \leq j_m \leq k$, such that $P^{(j_m)}(\,E_m) = 1$. Since the E_m's are disjoint the j_m's must be distinct. Therefore $n \leq k$.

Proposition 32. *Each $P^{(k)}$-cl. set contains a $P^{(k)}$-indecomp. set and intersects a $P^{(k)}$-max. indecomp. set. The number of distinct $P^{(k)}$-max. indecomp. sets is the maximum number of disjoint $P^{(k)}$-cl. sets.*

Proof. If there were a $P^{(k)}$-cl. set which does not contain any $P^{(k)}$-indecomp. subset then the set itself is $P^{(k)}$-decomp. and hence contains two disjoint $P^{(k)}$-cl. sets each of which is $P^{(k)}$-decomp. Hence by induction there would be an infinite number of disjoint $P^{(k)}$-cl. sets, contradicting Proposition 31. Now by the $P^{(k)}$-version of Proposition 15, each $P^{(k)}$-indecomp. set is contained in a $P^{(k)}$-max. indecomp. set; hence each $P^{(k)}$-cl. set intersects a $P^{(k)}$-max. indecomp. set. Two disjoint $P^{(k)}$-cl. sets cannot intersect the same $P^{(k)}$-max. indecomp. set, proving the last assertion.

Proposition 33. *For each k let $\delta(k)$ be the number of distinct $P^{(k)}$-max. indecomp. sets contained in X; then $\delta(k) \mid k$. These $\delta(k)$ sets form a clean cycle $\{I_i, 1 \leq i \leq \delta(k)\}$. $X - \bigcup_{i=1}^{\delta(k)} I_i$ does not contain any $P^{(k)}$-cl. sets and is not abs. ess.*

Proof. By Proposition 32, there exists a $P^{(k)}$-max. indecomp. set I. Set

$$I_i = \mathfrak{A}^i(I), \qquad i \geq 0.$$

By Proposition 28, $\{I_{k-i}, 1 \leq i \leq k\}$ is the k-cycle generated by I. We have $I = I_0 \subset I_k$. But by Proposition 29, each I_i is $P^{(k)}$-max. indecomp. Hence by the $P^{(k)}$-version of Proposition 16, $I_0 = I_k$ and consequently $I_i = I_j$ if $i \equiv j \pmod k$. Let d be the least positive integer such that $I_0 = I_d$. Then $I_i \neq I_j$ for $0 \leq i < j \leq$ $\leq d - 1$, for otherwise one would have

$$I_0 = I_k = \mathfrak{A}^{k-i}(I_i) = \mathfrak{A}^{k-i}(I_j) = \mathfrak{A}^{k+j-i}(I_0) = I_{k+j-i} = I_{j-i},$$

contradicting the definition of d. By the $P^{(k)}$-version of Proposition 16, the sets I_i, $1 \leq i \leq d$, are disjoint and so form a clean cycle. We have now $I_i = I_j$ if and only if $i \equiv j \pmod d$, hence $d \mid k$.

This d is the $\delta(k)$ asserted in the proposition, that is, any $P^{(k)}$-max. indecomp. set is one of the I_i's. To see this let J be such a set. As before, there is an integer e such that $e|k$ and $\{\mathfrak{A}^j(J),\ 1 \leqq j \leqq e\}$ is a cycle. Let $D = \bigcup_{j=1}^{e} \mathfrak{A}^j(J)$. Since X is indecomp., $C \cap D \neq 0$ and consequently $\mathfrak{A}^i(I) \cap \mathfrak{A}^j(J) \neq 0$ for some i and j. But then $\mathfrak{A}^i(I) = \mathfrak{A}^j(J)$ because both sets are $P^{(k)}$-max. indecomp. and it follows that

$$J = \mathfrak{A}^k(J) = \mathfrak{A}^{k+i-j}(I).$$

Thus J is one of the I_i's and therefore $d = \delta(k)$.

By Proposition 32, any $P^{(k)}$-cl. set must intersect one of the I_i's. Hence $X - C$ does not contain any $P^{(k)}$-cl. set. Then $X - C$ is not abs. ess. by the $P^{(k)}$-version of Proposition 14 and Proposition 25.

Definition 8. $\delta(k)$ is called *the cyclic index belonging to* k and the $\delta(k)$-cycle described in Proposition 33 is called the *cycle belonging to* k. It is uniquely defined for each k.

Notation. For two positive integers k and k' we denote their least common multiple by $k \vee k'$ and their greatest common divisor by $k \wedge k'$.

Proposition 34. *For arbitrary k and k', we have*

$$\delta(k \vee k') = \delta(k) \vee \delta(k') \tag{7}$$

$$\delta(k \wedge k') = \delta(k) \wedge \delta(k'). \tag{8}$$

Proof. Let $\{D_i,\ 1 \leqq i \leqq \delta(k)\}$ and $\{E_i,\ 1 \leqq i \leqq \delta(k')\}$ be the cycles belonging to k and k' respectively.

We first show that

$$\delta(k) \geqq k \wedge \delta(k'). \tag{9}$$

Writing $d = k \wedge \delta(k')$ and $\delta(k') = qd$, we set

$$F_r = \bigcup_{m=0}^{q-1} E_{md+r}.$$

The sets $\{F_r,\ 1 \leqq r \leqq d\}$, are clearly disjoint and $P^{(d)}$-cl., hence $P^{(k)}$-cl. It follows from Proposition 32 that there are at least d distinct $P^{(k)}$-max. indecomp. sets; hence $\delta(k) \geqq d$, which is (9).

Next, we show that

$$\text{if } k|k' \quad \text{then} \quad \delta(k)|\delta(k'). \tag{10}$$

Since X is indecomp., and $\bigcup_{i=1}^{\delta(k)} D_i$ and $\bigcup_{i=1}^{\delta(k')} E_i$ are both $P^{(1)}$-cl., we have $D_i \cap E_i \neq 0$ for some i and j. By relabelling we may suppose that $D_1 \cap E_1 \neq 0$. Define D_i and E_i for all $i \geqq 1$ by setting $D_i = D_j$ if $i \equiv j \pmod{\delta(k)}$ and $E_i = E_j$ if $i \equiv j \pmod{\delta(k')}$. Then it follows from the properties of cycles that the sets $D_i \cap E_i$, $1 \leqq i \leqq \delta(k) \vee \delta(k')$, are disjoint and $P^{(k \vee k')}$-cl., hence $P^{(k')}$-cl. if $k|k'$. Hence

$$\delta(k) \vee \delta(k') \leqq \delta(k')$$

by Proposition 32 and consequently (10) is true.

We can now prove that

$$\text{(11)} \qquad \text{if } k \,|\, k' \quad \text{then} \quad \delta(k) = k \wedge \delta(k').$$

For $\delta(k)\,|\,k$ by Proposition **33**; together with (10) this implies

$$\delta(k) \leqq k \wedge \delta(k').$$

Together with (9) this implies (11).

Let $k \vee k' = l$, then we have by (11):

$$\text{(12)} \qquad \delta(k) = k \wedge \delta(l), \quad \delta(k') = k' \wedge \delta(l).$$

Since $\delta(l)\,|\,l$ it is a simple arithmetical fact that

$$\text{(13)} \qquad (k \wedge \delta(l)) \vee (k' \wedge \delta(l)) = (k \vee k') \wedge \delta(l) = l \wedge \delta(l) = \delta(l).$$

Substituting from (12) into (13) we obtain (7).

Finally, let $k \wedge k' = d$; then it follows from (12) that

$$\text{(14)} \qquad \delta(d) = d \wedge (\delta(k) \wedge \delta(k')).$$

Since $\delta(k)\,|\,k$ and $\delta(k')\,|\,k'$ by Proposition 33, (14) reduces to (8).

Proposition 35. *We have for an arbitrary k,*

$$\text{(15)} \qquad \delta(\delta(k)) = \delta(k);$$

and the cycle belonging to $\delta(k)$ coincides, member for member, with that belonging to k.

Proof. Writing $d = \delta(k)$, we observe that each I_i in Proposition **33** is $P^{(d)}$-cl. and $P^{(k)}$-max. indecomp. Hence it is $P^{(d)}$-max. indecomp. by the last assertion in Proposition 30. Thus $\delta(d) \geqq d$ and since $\delta(d)\,|\,d$ we have $\delta(d) = d$. The rest follows.

The equation (15) also follows from (8) if we substitute $\delta(k)$ for k' there and use the fact that $\delta(k)\,|\,k$.

Proposition 36. *To each prime number p there corresponds an e_p which is either a nonnegative integer or „infinite", such that*

$$\delta(p^n) = p^{\min(n, e_p)}$$

for each $n \geqq 1$.

Proof. For each prime p define $e = e_p$ to be the least nonnegative integer such that $\delta(p^{e+1}) \neq p^{e+1}$, or ∞ if such an integer does not exist. Then $\delta(p^n) = p^n$ for $0 \leqq n < e + 1$. If $e = \infty$ there is nothing more to prove. Suppose now $0 \leqq e < \infty$, then by (10):

$$p^e = \delta(p^e) \,|\, \delta(p^{e+1}) < p^{e+1},$$

so that $\delta(p^{e+1}) = p^e$. Hence for each $n \geqq e + 1$ we have by (11):

$$p^e = \delta(p^{e+1}) = p^{e+1} \wedge \delta(p^n).$$

It follows that $\delta(p^n) = p^e$ since $\delta(p^n) \,|\, p^n$.

Proposition 37. *Let*
$$k = \prod_p p^{f_p}$$
be the prime-factorization of k, *then*
$$\delta(k) = \prod_p p^{\min(f_p, e_p)}$$
where e_p *is as given in Proposition 36.*

Proof. This is an immediate consequence of Proposition 36 and equation (7).

§ 4. Consequent sets

Definition 9. The set C in $\mathscr{B}$ is called a k^{th} *consequent of* x if $P^{(k)}(x, C) = 1$. The sequence $\{C_k, k \geqq 1\}$ is called a *consequent sequence of* x if for each $k \geqq 1$, C_k is a k^{th} consequent of x.

Proposition 38. *Given a consequent* $\{C_k, k \geqq 1\}$ *of* x, *there exists a consequent sequence* $\{D_k, k \geqq 1\}$ *of* x *such that* $D_k \subset C_k$ *and* $D_k \subset \mathfrak{A}(D_{k+1})$.

Proof. Let
$$D_k = \bigcap_{j=0}^{\infty} \mathfrak{A}^j(C_{k+j}).$$
Then $D_k \subset \mathfrak{A}^0(C_k) = C_k$; and
$$\mathfrak{A}(D_{k+1}) = \bigcap_{j=0}^{\infty} \mathfrak{A}^{j+1}(C_{k+1+j}) = \bigcap_{j=1}^{\infty} \mathfrak{A}^j(C_{k+j}) \supset D_k.$$
Since $P^{(k+j)}(x, C_{k+j}) = 1$ for each $j \geqq 0$, we have by Proposition 26,
$$P^{(k)}(x, \mathfrak{A}^j(C_{k+j})) = 1;$$
and consequently
$$P^{(k)}(x, \bigcap_{j=0}^{\infty} \mathfrak{A}^j(C_{k+j})) = 1.$$
This proves that D_k is a k^{th} consequent of x for each $k \geqq 1$.

Definition 10. For each x we define a probability measure $\pi_x(\cdot)$ as follows: for each $E \in \mathscr{B}$,
$$\pi_x(E) = \sum_{n=1}^{\infty} \frac{1}{2^n} P^{(n)}(x, E).$$
It is clear that $\pi_x(\cdot)$ is a probability measure, and that $\pi_x(E) = 0$ if and only if $L(x, E) = 0$, or equivalently if and only if $x \in E^0$.

Definition 11. A k^{th} consequent C of x is called *minimal* if C is minimal with respect to the measure π_x, namely if there does not exist a k^{th} consequent D with $\pi_x(D) < \pi_x(C)$. A *minimal consequent sequence* is one in which each member is minimal.

Proposition 39. *For each* x *and each consequent sequence* $\{C_k, k \geqq 1\}$ *of* x, *there exists a minimal consequent sequence* $\{D_k, k \geqq 1\}$ *such that* $D_k \subset C_k$ *for each* $k \geqq 1$.

Proof. There always exists a consequent sequence of x, namely the sequence all members of which are X. Writing for a moment $C \in \mathscr{C}_k(x)$ if C is a k^{th} consequent of x we set
$$a_k = \inf_{C \in \mathscr{C}_k(x)} \pi_x(C).$$

Then there exists a $C_{k,n}$ in $\mathscr{C}_k(x)$ with $\pi_x(C_{k,n}) < a_k + \frac{1}{n}$. Let $D_k = C_k \cap \bigcap_{n=1}^{\infty} C_{k,n}$; then D_k is a k^{th} consequent of x and $\pi_x(C_k) = a_k$. Clearly $\{D_k, k \geqq 1\}$ is a minimal consequent sequence and $D_k \subset C_k$ for each $k \geqq 1$.

Proposition 40. *In an indecomp. space X two minimal k^{th} consequents of a given x differ by a set which is not abs. ess.*

Let C_k and D_k be two minimal k^{th} consequents of x, then $\pi_x(C_k \triangle D_k) = 0$ and so by a previous remark $(C_k \triangle D_k)^0 \neq 0$. Consequently $C_k \triangle D_k$ is not abs. ess. by Proposition 17.

Proposition 41. *Let X be indecomp., x an arbitrary point of X, and $\{C_n, n \geqq 1\}$ an arbitrary consequent sequence of x. There exists a not abs. ess. set F (depending on x) and for each $y \in X - F$ there exists a positive integer $m(y)$ such that $\{C_{m(y)+n}, n \geqq 1\}$ is a consequent sequence of y.*

Proof. We have for each pair of integers m and n with $m < n$:

$$1 = P^{(n)}(x, C_n) = \int_X P^{(n-m)}(y, C_n)\, P^{(m)}(x, dy).$$

Hence there is a set $F_{m,n}$ in $\mathscr{B}$ with $P^{(m)}(x, F_{m,n}) = 0$ and such that if $y \in X - F_{m,n}$ then

$$P^{(n-m)}(y, C_n) = 1.$$

Let $F_m = \bigcup_{n=m+1}^{\infty} F_{m,n}$. Then $P^{(m)}(x, F_m) = 0$; and if $y \in X - F_m$, the above equation holds for every $n \geqq m+1$. Let $F = \bigcap_{m=1}^{\infty} F_m$, then $F \in \mathscr{B}$ and $P^{(m)}(x, F) = 0$ for every $m \geqq 1$. Consequently $F^0 \neq 0$ and F is not abs. ess. by Proposition 17. If $y \in X - F$, then there exists a positive integer $m = m(y)$ such that $y \in X - F_m$ and $P^{(k)}(y, C_{m+k}) = 1$ for every $k \geqq 1$. This proves the proposition.

In propositions 42 to 48 the space X is assumed to be abs. ess. and indecomp.

Proposition 42. *Let X be abs. ess. and indecomp. For each x there exists a finite positive integer $k(x)$ such that if $\{C_k, k \geqq 1\}$ is any consequent sequence of x then there exist m and n both less than $k(x) + 1$ such that $C_m \cap C_n$ is abs. ess.*

It is sufficient to prove this for a fixed minimal consequence $\{C_k, k \geqq 1\}$. For then the conclusion will remain valid with the same m and n for any consequent sequence of x by Proposition 40. Furthermore we may suppose on account of Propositions 39 and 38 that $C_k \subset \mathfrak{A}(C_{k+1})$. Hence $C = \bigcup_{k=1}^{\infty} C_k$ is cl. and consequently abs. ess. by Proposition 18. Set

$$D_k = C_k - \bigcup_{j=1}^{\infty}(C_k \cap C_{k+j}).$$

If $y \in D_k$ then $P^{(j)}(y, C_{k+j}) = 1$ and hence $P^{(j)}(y, D_k) = 0$ for each $j \geqq 1$, since $D_k \cap C_{k+j} = 0$. Thus $L(y, D_k) = 0$ and $D_k \subset D^0{}_k$. Such a D_k is clearly iness. and consequently $D = \bigcup_{k=1}^{\infty} D_k$ is not abs. ess. But

$$C - D = \bigcup_{k=1}^{\infty} \bigcup_{j=1}^{\infty} (C_k \cap C_{k+j}).$$

It follows that at least one $C_k \cap C_{k+j}$ is abs. ess., as was to be proved.

For each x let $\{C_k(x), k \geqq 1\}$ be a minimal consequent sequence of x and let

$$h(x) = \min\{|m-n| : C_m(x) \cap C_n(x) \text{ is abs. ess.}\},$$
$$h'(x) = \text{g. c. d. } \{|m-n| : C_m(x) \cap C_n(x) \text{ is abs. ess.}\},$$

where "g. c. d." stands for "the greatest common divisor". According to Proposition 42, $h'(x) \leqq h(x) \leqq k(x) < \infty$, and both $h(x)$ and $h'(x)$ are independent of the choice of the minimal consequent sequence.

Proposition 43. *There exist an integer H and a set F_H which is in $\mathscr{B}$ and not abs. ess. such that $h(x)$ is equal to H for all $x \in X - F_H$, and $h(x) \leqq H$ for all $x \in X$.*

Proof. Let x be an arbitrary point and let $\{C_n, n \geqq 1\}$ be a minimal consequent sequence of x such that $C_n \subset \mathfrak{U}^k(C_{n+k})$ for each $n \geqq 1$ and $k \geqq 1$. Such a choice is possible by Proposition 38. Let $h(x) = l$, then by definition there exists an integer j such that

$$C_j \cap C_{j+l} \text{ is abs. ess.}$$

For every y in this intersection, we have by the choice of $\{C_n, n \geqq 1\}$, for each $k \geqq 1$,

$$P^{(k)}(y, C_{j+k} \cap C_{j+k+l}) = 1.$$

Hence by Proposition 9,

$$C_{j+k} \cap C_{j+k+l} \text{ is abs. ess.,} \quad \text{for each } k \geqq 0,$$

or

$$C_n \cap C_{n+l} \text{ is abs. ess.,} \quad \text{for each } n \geqq j. \tag{16}$$

According to Proposition 41, there exists a set F in $\mathscr{B}$ which is not abs. ess. such that if $y \in X - F$, then $\{C_{m(y)+n}, n \geqq 1\}$ is a consequent sequence (but not necessarily minimal) of y for some $m(y) \geqq 1$. Hence we have $h(y) \geqq l$ by the definition of $h(\cdot)$.

We now prove that the function $h(\cdot)$ is bounded on X. For otherwise let $\{x_n, n \geqq 1\}$ be points of X such that $\lim\limits_{n\to\infty} h(x_n) = \infty$. By what we have proved, for each x_n there exists a set F_{x_n} in $\mathscr{B}$ which is not abs. ess. and such that

$$h(y) \geqq h(x_n) \quad \text{if} \quad y \in X - F_{x_n}.$$

Since X is abs. ess., $X - \bigcup\limits_{n=1}^{\infty} F_{x_n}$ is not empty; and if y is in this set, $h(y)$ would be ∞ which is impossible. Hence we may set

$$\max_{x \in X} h(x) = H < \infty.$$

By the argument above, there exists a not abs. ess. set F_H such that $h(x) \geqq H$, hence $h(x) = H$ on $X - F_H$, as was to be proved.

Remark. It has not been shown that the function h is $\mathscr{B}$-measurable, but this information will not be needed below.

Definition 12. The integer H is called the *overlapping index*, and the set $X - F_H$ (in $\mathscr{B}$) the *overlapping core* of the abs. ess. and indecomp. space X.

Proposition 43.1. *For each x in $X - F_H$ there exists an integer $\nu(x)$ such that for an arbitrary consequent sequence $\{C_n, n \geqq 1\}$ of x,*

$$C_n \cap C_{n+H} \textit{ is abs. ess. for } n \geqq \nu(x).$$

Proof. This is merely a restatement of (16).

Proposition 44. *For each k, $\delta(k) \mid H$.*

Proof. Consider the cycle $\{I_i, 1 \leqq i \leqq \delta(k)\}$ belonging to k and set $I_i = I_j$ if $i \equiv j (\mathrm{mod}\, \delta(k))$. Then $C = \bigcup_{i=1}^{\delta(k)} I_i$ is abs. ess. by Proposition 33, since X is abs. ess. and $X - C$ is not. If $x \in C \cap (X - F_H)$, then $\{I_i, i \geqq 1\}$ is a consequent sequence of x, and we have by Proposition 43,

$$I_i \cap I_{i+H} \neq 0$$

for some i. But the cycle is clean according to Proposition 33, hence $\delta(k) \mid H$.

Definition 13. Let $D = \max_{k \geqq 1} \delta(k)$; D is called the *maximum cyclic index* and the cycle belonging to D is called the *maximum cycle.*

Proposition 45. *In the notation of Proposition 36, we have*

$$D = \prod_p p^{e_p}$$

where $0 \leqq e_p < \infty$ for each prime p and also $e_p > 0$ for only a finite number of values of p. Furthermore, we have for each $k \geqq 1$,

$$\delta(k) = k \wedge D. \tag{17}$$

Proof. This is immediate from Propositions 36, 37 and 44, the last implying that $e_p < \infty$ for each p. A more direct proof of (17) is as follows. Let $\delta(k') = D$, then by (9),

$$\delta(k) \geqq k \wedge D. \tag{18}$$

On the other hand, by (7),

$$\delta(k \vee k') = \delta(k) \vee D;$$

hence $\delta(k) \mid D$ for otherwise one would have $\delta(k \vee k') > D$ which is impossible by the definition of D. Since $\delta(k) \mid k$ it follows that $\delta(k) \mid (k \wedge D)$ and so there must be equality in (18).

Example 1. $X = \{1, 2, 3, 4, 5\}$.

$$P(n, n+1) = 1 \quad \text{for} \quad n = 1, 2, 3;$$
$$P(4, 1) = 1; \quad P(5, 1) = P(5, 2) = \tfrac{1}{2}.$$

Each $\{n\}$, $n = 1, 2, 3, 4$, is $P^{(4)}$-max. indecomp.; $\{1, 3\}$ and $\{2, 4\}$ are $P^{(2)}$-indecomp., but $\{2, 4\}$ is not $P^{(2)}$-max. indecomp. since $\{2, 4, 5\}$ is. This example shows that the cycle belonging to a divisor of k is not necessarily obtained by the obvious grouping from the cycle belonging to k.

Example 2. $X = \{1, 2, 3, 4, 5, 6, 7, 8\}$.

$$P(1, 5) = P(1, 6) = \tfrac{1}{2};$$
$$P(2, 5) = P(2, 6) = P(2, 7) = P(2, 8) = \tfrac{1}{4};$$
$$P(3, 7) = P(4, 8) = P(5, 3) = P(6, 4) = P(7, 1) = P(8, 2) = 1.$$

Here the maximum index $D = 2$ and the maximum cycle is composed of $\{1, 2, 3, 4\}$ and $\{5, 6, 7, 8\}$. It is easily verified that $H = 2$ and $F_H = 0$.

The minimal consequent sequence for $\{6\}$ is

$$\{6\}, \{4\}, \{8\}, \{2\}, \{5, 6, 7, 8\}, \{1, 2, 3, 4\}, \{5, 6, 7, 8\}, \ldots .$$

If we denote this sequence of sets by $\{C_n, n \geqq 0\}$, it is to be noted that $C_1 \cap C_n = 0$ for $n = 2, 3, 4$ but $C_2 \cap C_4 \neq 0$.

Example 3. $X = \{1, 2, 3, 4, 5, 6, 7, 8\}$.

$$P(1, 2) = P(1, 4) = P(3, 4) = P(3, 6) = \tfrac{1}{2};$$
$$P(2, 3) = P(4, 5) = P(5, 6) = P(6, 7) = P(7, 8) = P(8, 1) = 1.$$

The minimal consequent sequence for $\{1\}$ is:

$$\{1\}, \{2, 4\}, \{3, 5\}, \{4, 6\}, \{5, 7\}, \{6, 8\}, \{7, 1\},$$
$$\{2, 4, 8\}, \{1, 3, 5\}, \{2, 4, 6\}, \{3, 5, 7\}, \{4, 6, 8\}, \{5, 7, 1\},$$
$$\{2, 4, 6, 8\}, \{1, 3, 5, 7\}, \{2, 4, 6, 8\}, \ldots .$$

Here in notation similar to the above: $C_1 \cap C_3 \neq 0$, $C_3 \cap C_5 \neq 0$, but $C_1 \cap C_5 = 0$.

Proposition 46. *There exist an integer H' and a set $F'_{H'}$ which is in $\mathscr{B}$ and not abs. ess. such that $h'(x)$ is equal to H' for all $x \in X - F'_{H'}$, and $h'(x) \leqq H'$ for all $x \in X$.*

Proof. Let

$$H' = \max_{x \in X} h'(x);$$

since $h'(x) \leqq h(x)$ for every x, we have $H' \leqq H < \infty$. The rest of the proof is exactly the same as the first part of the proof of Proposition 43.

Remark. In Propositions 43 and 46, we may replace the sets $X - F_H$ and $X - F'_{H'}$ by cl. subsets. For if we set

$$G = (X - F_H) \cap (X - F_H)^0, \tag{19}$$

$$G' = (X - F'_{H'}) \cap (X - F'_{H'})^0, \tag{20}$$

then G and G' are cl. by Proposition 1, and $X - G$ and $X - G'$ are not abs. ess. by Proposition 18.1.

The next proposition is due to S. T. C. Moy.

Proposition 47. $D = H'$.

Proof. Choose any x in $K \cap G'$ where K is the maximum cycle and G' is given in (20), $K \cap G'$ being cl. by Proposition 18. Let $\{C_j(x), j \geqq 1\}$ be a minimal consequent sequence of x. Then $C_m(x) \cap C_n(x) \neq 0$ implies $D \mid (m - n)$; hence $D \mid H'$. On the other hand, let us set for such an x:

$$E_r = \bigcup_{n=0}^{\infty} C_{nH'+r}(x), \quad 0 \leqq r \leqq H' - 1$$

where the C_j's have been chosen to satisfy

$$C_j(x) \subset \mathfrak{A}(C_{j+1}(x)) \tag{21}$$

by Proposition 38. Then each E_r is $P^{(H')}$-cl. and the H' sets $\{E_r, 0 \leqq r \leqq H' - 1\}$ form a H'-cycle. If F denotes the union of the pairwise intersections of the E_r's, F is not abs. ess. by the definition of $H' = h'(x)$. Hence $F^0 \cap F^c$ is cl. by Propositions 19 and 1. The H' sets $F^0 \cap F^c \cap E_r$ are disjoint; their union is nonempty

since the union of the E_r's is cl. and X is indecomp., hence each $F^0 \cap F^c \cap E_r$ is nonempty and so $P^{(H')}$-cl. by properties of a cycle. Thus we have by Propositions **32** and **33**:

$$D \geqq \delta(H') = H'.$$

Hence $D = H'$.

The next proposition, conjectured by the author, was first proved by H. KESTEN. The version given below, using an essential idea of his, is simpler.

Proposition 48. $H' = H$.

Lemma 1. *Let x be arbitrary and $\{C_n, n \geqq 1\}$ a minimal consequent sequence of x. Suppose that $n \neq m$ and*

$$C_m \cap C_n \text{ is abs. ess.,} \tag{22}$$

then

$$P^{(n)}(x, C_m) > 0. \tag{23}$$

Proof. By Proposition 19, (22) implies that $(C_m \cap C_n)^0 = 0$. It follows that

$$\pi_x(C_m \cap C_n) > 0,$$

and consequently

$$\pi_x(C_n \setminus C_m) < \pi_x(C_n).$$

Since C_n is a minimal n^{th} consequent set of x, $C_n \setminus C_m$ cannot be likewise. Thus

$$P^{(n)}(x, C_n \setminus C_m) < 1$$

which implies (23).

Lemma 2. *Let the hypotheses in Lemma 1 hold for an x in G where G is given by* (19). *Then there exists an $l \geqq 1$ such that*

$$C_{n+H+l} \cap C_{m+l} \text{ is abs. ess.} \tag{24}$$

Proof. We may choose the C_j's to satisfy (21) and furthermore $C_j \subset G$ for every $j \geqq 1$. Let $y \in C_n$ and $\{D_j(y), j \geqq 1\}$ be a minimal consequent sequence of y. Owing to (21) we may suppose that for every y in C_n we have $D_j(y) \subset C_{j+n}$ for every $j \geqq 1$. Since $h(y) = H$, there exists a $j = j(y)$ such that

$$D_j(y) \cap D_{H+j}(y) \text{ is abs. ess.}$$

It follows from Lemma 1 that

$$P^{(j)}(y, C_{n+H+j}) \geqq P^{(j)}(y, D_{H+j}(y)) > 0. \tag{25}$$

If $j < k$, we have by (21)

$$P^{(j)}(y, C_{n+H+j}) \leqq P^{(k)}(y, C_{n+H+k}).$$

Consequently if we set

$$C_{n,k} = \left\{y \in C_n : P^{(k)}(y, C_{n+H+k}) \geqq \frac{1}{k}\right\},$$

then $C_n = \bigcup\limits_{k=1}^{\infty} C_{n,k}$ and in particular

$$C_m \cap C_n = \bigcup_{k=1}^{\infty} [C_m \cap C_{n,k}].$$

The hypothesis (22) then implies the existence of an l such that

$$E = C_m \cap C_{n,l} \text{ is abs. ess.}$$

Let us write also

$$F = C_{m+l} \cap C_{n+H+l}.$$

For each $y \in E$, it follows from (21) and the definition of C_{n+l} that

$$P^{(l)}(y, F) = P^{(l)}(y, C_{n+H+l}) \geqq \frac{1}{l}.$$

Therefore

$$\inf_{x \in E} L(x, F) \geqq \frac{1}{l} > 0.$$

Since E is abs. ess., this implies F is abs. ess. by Proposition 9.

Lemma 3. *Under the same hypotheses as in Lemma 2, we have* $H \mid (m - n)$.

Proof. We may suppose that $n < m$ and

$$m - n = qH + r, \quad 0 \leqq q, \quad 1 \leqq r \leqq H.$$

Applying Lemma 2 successively q times, we infer that there exists an $l \geqq 1$ such that

$$C_{n+qH+l} \cap C_{m+l} \text{ is abs. ess.}$$

By the definition of $H = h(x)$, this implies

$$H \leqq (m + l) - (n + qH + l) = r \leqq H$$

Hence $r = H$ and $H \mid (m - n)$.

Proof of Proposition 48. Choose any x in $G \cap G'$ (see (19) and (20)) which is nonempty since the space is indecomp. Then $h(x) = H$, $h'(x) = H'$. By the definitions, we have $h'(x) \leqq H$. Furthermore it follows from Lemma 3 that $H \mid h'(x)$. Hence $H' = h'(x) = H$.

Proposition 48.1. *For every* $x \in G \cap G'$, *we have*

$$h(x) = h'(x) = H = H' = D.$$

§ 5. Decomposition theorems

Proposition 49. *Suppose that* X *is indecomp. and abs. ess. For each* x *in* X *there exists a cl. set* C *such that: if* $E \subset C$ *then either* E *is abs. ess. or* $Q(x, E) = 0$.

Proof. Let $\mathscr{C}$ be the family of cl. sets in X, and set

$$\alpha = \alpha(x) = \inf_{C \in \mathscr{C}} L(x, C). \tag{26}$$

For each n, there exists a cl. set C_n such that

$$L(x, C_n) \leqq \alpha + \frac{1}{n}.$$

Let $C = C(x) = \bigcap_{n=1}^{\infty} C_n$. We have

$$L(x, C) \leqq \lim_{n \to \infty} L(x, C_n) \leqq \alpha;$$

hence

$$L(x, C) = \alpha$$

by the definition of α. Furthermore $\alpha > 0$ since $C^0 = 0$.

If $E \subset C$ and E is not abs. ess., then either E is iness. and so $Q(y, E) = 0$ for every $y \in X$; or E is imp. ess. In the latter case $E^0 \neq 0$ by Proposition 19, and $E^0 \cap C$ is cl. by indecomposability. Starting from x, if the process $\{\xi_n, n \geqq 0\}$ is in E infinitely often, then it must be in C infinitely often and never in $E^0 \cap C$; it follows that

$$Q(x, E) \leqq L(x, C) - L(x, E^0 \cap C).$$

Both terms on the right side are equal to α by the definition of α and C, hence $Q(x, E) = 0$ as was to be proved.

Let us write, for any x in X and E in $\mathscr{B}$:

$$M(x, E) = 1 - Q(x, E^c). \tag{27}$$

Thus $M(x, E)$ is the probability that the process starting from x ultimately stays in E. In this notation the set E is perp. (Definition 4) if and only if $M(x, E) > 0$ for some x in X.

Proposition 49.1. *Let $\mathscr{N}$ be the family of all sets which are not abs. ess., and $\alpha(x)$ be defined as in* (26), *then for each x:*

$$\sup_{E \in \mathscr{N}} M(x, E) = 1 - \alpha(x).$$

Proposition 50. *Let X be arbitrary, C a cl. subset such that $X - C$ does not contain any cl. set. Then there exists a sequence of disjoint iness. (possibly empty) sets $\{E_i, i \geqq 1\}$ such that*

$$X - C = \bigcup_{i=1}^{\infty} E_i; \tag{28}$$

$$\lim_{n \to \infty} P^{(n)}(x, C) = L(x, C) \tag{29}$$

for each x; and

$$\lim_{n \to \infty} P^{(n)}(x, \bigcup_{i=j+1}^{\infty} E_i) = 1 - L(x, C) \tag{30}$$

for each $j \geqq 0$.

Proof. Let

$$E_i = \left\{x \in X - C : \frac{1}{i} \leqq L(x, C) < \frac{1}{i-1}\right\}$$

for $i \geqq 1$. Since $X - C$ does not contain any cl. set, $C^0 = 0$ and consequently (28) holds. The set E_1 is clearly iness., and each E_i, $i \geqq 2$, is iness. by Proposition 10. Since C is cl., we have

$$\sum_{\nu=1}^{n} K^{(\nu)}(x, C) \leqq P^{(n)}(x, C) \leqq \sum_{\nu=1}^{\infty} K^{(\nu)}(x, C).$$

Letting $n \to \infty$ we obtain (29). Furthermore we have for arbitrary x in X and E in $\mathscr{B}$:

$$\overline{\lim_{n \to \infty}} P^{(n)}(x, E) \leqq Q(x, E).$$

Since the union of a finite number of iness. sets is iness., it follows that

$$\overline{\lim_{n \to \infty}} P^{(n)}(x, \bigcup_{i=1}^{j} E_i) \leqq Q(x, \bigcup_{i=1}^{j} E_i) = 0. \tag{31}$$

Finally we have

$$(32)\qquad 1 = P^{(n)}(x, X) = P^{(n)}(x, C) + P^{(n)}(x, \bigcup_{i=1}^{j} E_i) + P^{(n)}(x, \bigcup_{i=j+1}^{\infty} E_i)$$

Hence (30) follows from (29), (31) and (32).

Proposition 51. *Let X be arbitrary and φ be a σ-finite measure on $(X, \mathscr{B})$ such that if C is cl. then $\varphi(C) > 0$. Then there exists a set A which is the union of at most a denumerable number of indecomp. sets and such that A^0 does not contain any indecomp. set and is imp. ess. Furthermore $X - A - A^0$ does not contain any cl. set and is not abs. ess.*

Proof. It is well known that from a σ-finite measure one can construct a finite measure which is co-positive, hence we may suppose φ to be finite. Let the family of all indecomp. sets be $\{B_\alpha\}$ and let $A_\alpha = (B_\alpha^0)^0$. Each A_α is max. indecomp. by Proposition 15. Since $\varphi(A_\alpha) > 0$ for each α and $\varphi(X) < \infty$, the family of distinct A_α's is at most denumerable by Proposition 16. We put

$$A = \bigcup_\alpha A_\alpha .$$

The set A^0 is either cl. or empty, and since it is disjoint from A it cannot contain any indecomp. set by the definition of A. By Proposition 14, $X - A - A^0$ does not contain any cl. set and is not abs. ess. It remains to prove that A^0 is imp. ess. if not empty.

The following proof, considerably shorter than DOEBLIN's (cf. my Columbia lecture notes), is due to T. E. HARRIS.

Let $\varphi(A^0) = \lambda > 0$. For each x in A^0 let $\mathscr{C}(x)$ be the family of cl. sets containing x and let

$$\tilde{\varphi}(x) = \inf_{C \in \mathscr{C}(x)} \varphi(C) .$$

Observing that any sequence of sets in $\mathscr{C}(x)$ has a cl. intersection since it is nonempty, we deduce by the usual argument the existence of a set C_x in $\mathscr{C}(x)$ such that

$$\tilde{\varphi}(x) = \varphi(C_x) > 0 .$$

For each $n \geqq 1$, let

$$E_n = \left\{x \in A^0 : \tilde{\varphi}(x) \leqq \frac{\lambda}{n}\right\} .$$

If $y \in C_x$, then $\tilde{\varphi}(y) \leqq \varphi(C_x)$. It follows that if $x \in E_n$, then $C_x \subset E_n$ so that E_n is cl. for each $n \geqq 1$. Furthermore $A^0 = E_1 \supset E_2 \supset \ldots$, and $\bigcap_n E_n = 0$. For otherwise $\bigcap_n E_n$ would be cl. and if y were any point in it, $\tilde{\varphi}(y)$ would be zero which is impossible. We have therefore

$$(33)\qquad A^0 = \bigcup_n (A^0 - E_n) .$$

Suppose $A^0 - E_n$ were to contain a cl. set, then it would contain a cl. set with arbitrarily small φ-measure since every cl. subset of A^0 is decomp. In particular it would contain a cl. set F with $\varphi(F) \leqq \lambda/n$. Let $y \in F$, then

$$\tilde{\varphi}(y) \leqq \varphi(F) \leqq \frac{\lambda}{n} ,$$

which is impossible since $y \in A^0 - E_n$ implies $\tilde{\varphi}(y) > \lambda/n$. Hence for each n, $A^0 - E_n$ does not contain any cl. set and so is not abs. ess. by Proposition 11. It follows that A^0 is not abs. ess. by (33); but since A^0 is cl. it is ess. Thus A^0 is imp. ess. as was to be proved.

Proposition 52. *Let X be indecomp. and abs. ess. and φ be a σ-finite measure on $(X, \mathscr{B})$ such that if A is perp. then $\varphi(A) > 0$. Then we have*

$$X = B \cup C, \quad B \cap C = 0; \tag{34}$$

where B is perp. and imp. ess., C is cl. and every ess. subset E of C is abs. ess. and satisfies the relation

$$C \subset E^\infty. \tag{35}$$

Proof. As in the proof of Proposition 51 we may suppose that φ is a finite measure. Let $\mathscr{S}$ be the family of perp. and imp. ess. sets and let

$$\alpha = \sup_{A \in \mathscr{S}} \varphi(A).$$

We deduce by the usual argument the existence of a set A in $\mathscr{S}$ such that $\varphi(A) = \alpha$. Clearly $X - A$ does not contain any set in $\mathscr{S}$. Now take

$$B = (X - A^0) \cup A, \quad C = X - B = A^0 \cap (X - A).$$

Since $X - A^0$ is not abs. ess. by Proposition 18.1 we have $B \in \mathscr{S}$; C is cl. by Proposition 1. Since C does not contain any set in $\mathscr{S}$, it does not contain any imp. ess. set by Proposition 23.1. Hence any ess. subset E of C is abs. ess. By Proposition 19, E^∞ is cl. and $E^0 = 0$. It follows from Proposition 14.1 that $C - CE^\infty$ is not abs. ess. so it is iness. by what has just been proved. Thus if $x \in C - CE^\infty$, we have by Proposition 7 and the inequality (6):

$$Q(x, E) \geqq Q(x, CE^\infty) \geqq 1 - Q(x, C - CE^\infty) = 1.$$

On the other hand, $Q(x, E) = 1$ if $x \in E^\infty$. Thus $Q(x, E) = 1$ for every $x \in C$, and this is equivalent to (35).

Bibliography

[1] Doeblin, W.: Éléments d'une théorie générale des chaînes simples constantes de Markoff. Ann. Sci. École Norm. Sup. (3), **57**, No. 2, 61–111 (1940).

[2] — Sur les propriétés asymptotiques de mouvement régis par certains types de chaînes simples. Bull. Math. Soc. Roum. Sci. **39**, No. 1, 57–115; No. 2, 3–61 (1937).

[3] Doob, J. L.: Stochastic Processes. New York: Wiley & Sons, 1953.

[4] Chung, K. L.: Markov Chains With Stationary Transition Probabilities. Berlin-Göttingen-Heidelberg: Springer 1960.

[5] Blackwell, D.: The existence of anormal chains. Amer. math. Soc. **51** 465–468 (1945)

Sequoia Hall
Stanford University
Stanford, California (USA)

(Received September 7, 1963)

FIELDS, OPTIONALITY AND MEASURABILITY.*

By K. L. CHUNG and J. L. DOOB.

1. Basic notions. Throughout this paper the following notation will be used:

$$T^0 = (0,\infty), \qquad T = (-\infty, +\infty), \qquad T^* = [-\infty, +\infty],$$
$$\forall t \in T^*: \qquad T_t = (-\infty, t].$$

The usual Borel field on T, trivially extended to T^* if need be, is denoted by $\mathcal{B}$. Its restriction to T_t is denoted by $\mathcal{B}_t$.

N is the set of all integers; R is the set of all rational numbers; the restrictions of N and R to T^0 are denoted by N^0 and R^0 respectively. Where it is not specified, the letters s, t, u denote elements of T^0 or T; the letters k, m, n elements of N^0 or N; depending on the context. The quantifier "$\forall t$" or "$\forall n$" will be omitted sometimes.

For two numbers or two numerical functions α, β with the same domain, we write

$$\alpha \wedge \beta = \min(\alpha, \beta), \qquad \alpha \vee \beta = \max(\alpha, \beta).$$

The lattice notation $\wedge$ and $\vee$ will also be used for Borel fields. In this case if $\{\mathcal{F}_\iota\}$ is any indexed family of Borel fields on the same set, we put

$$\bigwedge_\iota \mathcal{F}_\iota = \text{the largest Borel field contained in every } \mathcal{F}_\iota;$$

$$\bigvee_\iota \mathcal{F}_\iota = \text{the smallest Borel field containing every } \mathcal{F}_\iota.$$

If Ω is an abstract set (space), a *Borel field* (B. F.) on Ω is a collection of subsets of Ω which is closed under complementation in Ω and countable union (and intersection). If $\Delta \subset \Omega$, we denote by

$$\Delta \cap \mathcal{F}$$

the collection of sets of the form $\Delta \cap F$ with F ranging over $\mathcal{F}$; this is seen to be a B. F. on Δ if Δ is not empty, otherwise it consists of the empty set.

If S is a subset of T^*, a family of B. F.'s $\{\mathcal{F}_s\}$ on Ω, indexed by S, is called nondecreasing iff

(1) $$\forall s < t: \qquad \mathcal{F}_s \subset \mathcal{F}_t.$$

Received June 8, 1964.

* This research is supported in part by the Office of Scientific Research of the United States Air Force.

Such a family can be trivially enlarged to one on the index set T^*, and we may suppose this to have been done in what follows. The following notation will be used:

$$\mathcal{F}_{t-} = \bigvee_{s<t} \mathcal{F}_s, \qquad \mathcal{F}_{t+} = \bigwedge_{s>t} \mathcal{F}_s.$$

Let α be a function on Ω to T^*. We shall write

$$\alpha \in \mathcal{F}$$

iff α is measurable with respect to the B. F. $\mathcal{F}$ in the usual sense, and say "α belongs to (or is contained in) $\mathcal{F}$" and "$\mathcal{F}$ contains α." This simplification of language seems overdue. The smallest B. F. containing a collection of functions such as $\{x_s, s \in S\}$ is also said to be *generated by* it (or them) and denoted by $\mathcal{F}\{x_s, s \in S\}$. In case of a single function, say α, the notation becomes $\mathcal{F}\{\alpha\}$. These definitions have their obvious generalizations if the range of the functions is in an abstract space, as will be supposed in §§ 3-4.

From now on, a B. F. is on Ω and a function is on Ω to T^*, unless otherwise specified. The set $\{\omega: \alpha(\omega) > t\}$, e. g., will be abbreviated as $\{\alpha > t\}$.

Definition 1. Let $\{\mathcal{F}_t, t \in T\}$ be a nondecreasing family of B. F.'s and α a function. The B. F. generated by the collections:

$$\{\alpha > t\} \cap \mathcal{F}_t, \qquad t \in T, \tag{2}$$

will be denoted by $\mathcal{F}_{\alpha-}$.

PROPOSITION 1. *An equivalent definition of $\mathcal{F}_{\alpha-}$ is obtained if we replace $\mathcal{F}_t$ in* (2) *by $\mathcal{F}_{t-}$ or $\mathcal{F}_{t+}$; or if we replace* (2) *by*

$$\{\alpha \geqq t\} \cap \mathcal{F}_{t-}, \qquad t \in T. \tag{2'}$$

Finally we may replace T in (2) *or* (2′) *by any dense subset of T.*

Proof. In this proof let us denote the B. F. obtained with $\mathcal{F}_{t-}$, $\mathcal{F}_t$, $\mathcal{F}_{t+}$ in (2) respectively by $\mathcal{B}_1$, $\mathcal{B}_2$, $\mathcal{B}_3$; and the B. F. generated by the sets in (2′) by $\mathcal{B}_4$. It follows from (1) that

$$\mathcal{B}_1 \subset \mathcal{B}_2 \subset \mathcal{B}_3.$$

On the other hand, we have for each $\Lambda \in \mathcal{F}_{t+}$,

$$\{\alpha > t\} \cap \Lambda = \bigcup_{n=1}^{\infty} [\{\alpha > t + 1/n\} \cap \Lambda]. \tag{3}$$

Clearly each member of the union above belongs to $\mathcal{B}_1$. Hence $\mathcal{B}_3 \subset \mathcal{B}_1$ and we have proved the equality of $\mathcal{B}_1$, $\mathcal{B}_2$ and $\mathcal{B}_3$.

Next, the equation (3) remains valid if the "$>$" on the right side is replaced by "$\geqq$," and now each member of the union there belongs to $\mathcal{B}_4$; hence $\mathcal{B}_3 \subset \mathcal{B}_4$. Conversely, for each m and $\Lambda \in \mathcal{F}_{t-(1/m)}$:

$$\{\alpha \geqq t\} \cap \Lambda = \bigcap_{n=m}^{\infty} [\{\alpha > t - (1/n)\} \cap \Lambda]$$

where each member of the intersection belongs to $\mathcal{B}_3$. Now it follows from a known result (see e. g. [5; p. 25]) that if $\Delta \subset \Omega$ and $\{\mathcal{F}_i\}$ is any collection of B. F.'s, we have

$$\bigvee_i (\Delta \cap \mathcal{F}_i) = \Delta \cap (\bigvee_i \mathcal{F}_i). \tag{4}$$

Hence $\forall t$:

$$\bigvee_{m=1}^{\infty} (\{\alpha \geqq t\} \cap \mathcal{F}_{t-(1/m)}) = \{\alpha \geqq t\} \cap \mathcal{F}_{t-}.$$

Consequently $\mathcal{B}_4 \subset \mathcal{B}_3$ and we have proved the equality of $\mathcal{B}_3$ and $\mathcal{B}_4$.

Finally, the last assertion of Proposition 1 follows from the equation: for each $\Lambda \in \mathcal{B}_t$,

$$\{\alpha > t\} \cap \Lambda = \bigcup_{t<r\in R} [\{\alpha > r\} \cap \Lambda]$$

if R is a dense subset of T.

PROPOSITION 2. *We have*

$$\mathcal{F}\{\alpha\} \subset \mathcal{F}_{\alpha-} \subset \mathcal{F}\{\alpha\} \vee \mathcal{F}_{+\infty}. \tag{5}$$

For each $B \in (t,\infty) \cap \mathcal{B}$, we have

$$\{\alpha \in B\} \cap \mathcal{F}_t \subset \mathcal{F}_{\alpha-}. \tag{6}$$

Proof. The first assertion is trivial but observe that we do not assume $\alpha \in \mathcal{F}_{+\infty}$. To prove (6) we note that

$$\forall u > t\colon \{\alpha > u\} \cap \mathcal{F}_t \subset \{\alpha > u\} \cap \mathcal{F}_u \subset \mathcal{F}_{\alpha-}.$$

Thus (6) is true if $B = (u,\infty)$ with $u > t$; hence it is true as asserted since the collection of sets for which it is true forms a Borel field on (t,∞).

PROPOSITION 3. *If $\alpha \leqq \beta$, then $\mathcal{F}_{\alpha-} \subset \mathcal{F}_{\beta-}$ if and only if $\alpha \in \mathcal{F}_{\beta-}$. If $\alpha_n \uparrow \alpha$*[1] *and $\forall n\colon \alpha_n \in \mathcal{F}_{\alpha-}$, then*

$$\bigvee_n \mathcal{F}_{\alpha_n-} = \mathcal{F}_{\alpha-} \tag{7}$$

[1] The symbols $\uparrow$ and $\downarrow$ are used for monotone convergence in the non-strict sense.

Proof. We have if $\alpha \leqq \beta$:

$$\forall t\colon \{\alpha > t\} \cap \mathcal{F}_t = \{\alpha > t\} \cap [\{\beta > t\} \cap \mathcal{F}_t]. \tag{8}$$

If $\alpha \in \mathcal{F}_{\beta-}$, then $\{\alpha > t\} \in \mathcal{F}_{\beta-}$ and the set on the right side above belongs to $\mathcal{F}_{\beta-}$. This proves $\mathcal{F}_{\alpha-} \subset \mathcal{F}_{\beta-}$. Conversely if the last inclusion holds, then by (5): $\alpha \in \mathcal{F}_{\alpha-} \subset \mathcal{F}_{\beta-}$. The first assertion of Proposition 3 is proved. It follows that we have "$\subset$" instead of "$=$" in (7). On the other hand, we have for each t and $\Lambda \in \mathcal{F}_t$:

$$\{\alpha > t\} \cap \Lambda = \bigcup_n [\{\alpha_n > t\} \cap \Lambda] \subset \bigvee_n \mathcal{F}_{\alpha_n-}.$$

Hence (7) is proved.

PROPOSITION 3.1. *If $\alpha_n \uparrow \infty$ and if $\forall n\colon \alpha_n \in \mathcal{F}_{+\infty}$, then*

$$\bigvee_n \mathcal{F}_{\alpha_n-} = \mathcal{F}_{+\infty}.$$

Definition 2. We define

$$\mathcal{F}_{\alpha+} = \bigwedge_{\delta>0} \mathcal{F}_{(\alpha+\delta)-}, \tag{9}$$

where the convention $\pm\infty + \delta = \pm\infty$ is used. It is obvious, even without the use of Proposition 3, that $\mathcal{F}_{(\alpha+\delta)-}$ is monotone in δ. Clearly $\mathcal{F}_{\alpha-} \subset \mathcal{F}_{\alpha+}$.

Remark. If we define $\mathcal{F}_{\alpha^*}$ to be the B. F. generated by the collections

$$\{\alpha \geqq t\} \cap \mathcal{F}_t, \qquad t \in T,$$

it is easy to see that

$$\mathcal{F}_{\alpha-} \subset \mathcal{F}_{\alpha^*} \subset \mathcal{F}_{\alpha+} = \bigwedge_{\delta>0} \mathcal{F}_{(\alpha+\delta)^*}.$$

We shall not use the intermediate B. F. $\mathcal{F}_{\alpha^*}$ except to note that if α is a constant t_0, then $\mathcal{F}_{\alpha-}$, $\mathcal{F}_{\alpha^*}$, $\mathcal{F}_{\alpha+}$ reduce respectively to $\mathcal{F}_{t_0-}$, $\mathcal{F}_{t_0}$, $\mathcal{F}_{t_0+}$. There are simple examples in which these three B. F.'s are in strictly increasing order.

Up to now the function α is arbitrary. We shall now turn our attention to a specially interesting class of functions.

Definition 3. The function α is called *optional relative to the non-decreasing family* $\{\mathcal{F}_t, t \in T\}$ of B. F.'s iff

$$\forall t \in T\colon \{\alpha < t\} \in \mathcal{F}_t; \tag{10}$$

it is called *strictly optional* iff

$$\forall t \in T\colon \{\alpha \leqq t\} \in \mathcal{F}_t. \tag{11}$$

If $\alpha \geqq 0$, then the index set T in the above may be replaced by T^0.

PROPOSITION 5. *An equivalent definition of optionality is obtained if we replace $\mathcal{F}_t$ in* (10) *by $\mathcal{F}_{t-}$ or $\mathcal{F}_{t+}$; or if we replace $\mathcal{F}_t$ in* (11) *by $\mathcal{F}_{t+}$; or if we replace T by any dense subset of T.*

The proof is similar to that of Proposition 1.

It follows from the definition that $\alpha \in \mathcal{F}_{+\infty}$, and that

$$\forall t \in T: \alpha^{-1}(\mathcal{B}_t) \subset \mathcal{F}_{t+} \text{ or } \mathcal{F}_t \tag{12}$$

according as α is optional or strictly optional. The next proposition is also trivial.

PROPOSITION 6. *Strict optionality implies optionality and the two notions coincide for a given $\{\mathcal{F}_t\}$ if and only if*

$$\forall t \in T: \mathcal{F}_t = \mathcal{F}_{t+}. \tag{13}$$

Furthermore, optionality relative to $\{\mathcal{F}_t\}$ is equivalent to strict optionality relative to $\{\mathcal{F}_{t+}\}$.

Unless otherwise specified, we shall regard the nondecreasing family $\{\mathcal{F}_t\}$ as given and optionality as relative to it. Clearly a constant function is strictly optional. The next two propositions, as well as Proposition 15 later, give ways of deriving new optional functions from given ones.

PROPOSITION 7. *If α and β are both (strictly) optional, then so is $\alpha \wedge \beta$ and $\alpha \vee \beta$; similarly for a finite number of terms. If $\forall n: \alpha_n$ is optional, then so is each one of the following:*

$$\sup_n \alpha_n, \quad \inf_n \alpha_n, \quad \limsup_n \alpha_n, \quad \liminf_n \alpha_n. \tag{14}$$

If $\forall n: \alpha_n$ is strictly optional, then so is the first one in (14), *the others being optional but not necessarily strictly so.*

Proof. It is convenient to use one of the equivalent conditions given in Proposition 5. For instance, since

$$\{\sup_n \alpha_n \leqq t\} = \bigcap_n \{\alpha_n \leqq t\}, \qquad \{\inf_n \alpha_n < t\} = \bigcup_n \{\alpha_n < t\},$$

the assertions follow quickly by applying the appropriate criteria. As regards the last assertion of the proposition, we may take any α which is optional but not strictly so (Example 1 in § 5), and $\alpha_n = \alpha + 1/n$. Then α_n is strictly optional and

$$\alpha = \inf_n \alpha_n = \lim_n \alpha_n.$$

PROPOSITION 8. *If α and β are both optional and nonnegative, then $\alpha+\beta$ is optional. If furthermore one of the following three conditions is satisfied, then $\alpha+\beta$ is strictly optional:*

(i) $\alpha>0, \beta>0$;

(ii) $\beta>0$, *β is strictly optional;*

(iii) *α and β are both strictly optional.*

Remark. The condition "$\alpha>0$ and β is strictly optional" is not sufficient. Take α to be positive, optional but not strictly so; and take $\beta=0$.

Proof. Let α and β be optional and nonnegative throughout this proof. For each $t>0$:

$$\{\alpha+\beta<t\}=\bigcup_{r\in R\cap(0,t)}\{\alpha<r;\beta<t-r\}\in\mathfrak{F}_t,$$

proving the first assertion. Next, consider the decomposition for $t\geqq 0$:

$$\begin{aligned}\{\alpha+\beta>t\} &= \{0<\alpha<t;\alpha+\beta>t\}\cup\{\alpha=0;\beta>t\}\\ &\qquad\cup\{\alpha>t;\beta=0\}\cup\{\alpha\geqq t;\beta>0\}.\end{aligned}$$

Let the four sets on the right side be denoted by Λ_1, Λ_2, Λ_3 and Λ_4 respectively. It is easy to see that for each $t>0$:

$$\Lambda_1=\bigcup_{r\in R\cap(0,t)}\{r<\alpha<t;\beta>t-r\}.$$

For each r in $(0,t)$ the set in $\{\ \}$ above belongs to $\mathfrak{F}_t$, hence $\Lambda_1\in\mathfrak{F}_t$.

Under (i), we have $\Lambda_2=0$, $\Lambda_3=0$, and if $t>0$:

$$\Lambda_4=\{\alpha\geqq t\}=\Omega-\{\alpha<t\}\in\mathfrak{F}_t.$$

Under (ii), we have $\Lambda_3=0$, and $\Lambda_4\in\mathfrak{F}_t$ as before. Furthermore if $t>0$:

$$\Lambda_2=\{\alpha\geqq 0\}\cap\{\beta>t\}\in\mathfrak{F}_{0+}\vee\mathfrak{F}_t=\mathfrak{F}_t.$$

Under (iii), we have if $t\geqq 0$:

$$\Lambda_2\in\mathfrak{F}_0\vee\mathfrak{F}_t=\mathfrak{F}_t,$$

and by symmetry $\Lambda_3\in\mathfrak{F}_t$. Furthermore

$$\Lambda_4=\{\alpha\geqq t\}\cap\{\beta>0\}\in\mathfrak{F}_t\vee\mathfrak{F}_0=\mathfrak{F}_t.$$

Hence, in cases (i) and (ii) we have

$$\{\alpha+\beta>t\}\in\mathfrak{F}_t$$

for every $t > 0$, while in case (iii) the same is true for every $t \geqq 0$. Proposition 8 is completely proved.

It should be observed that Proposition 8 is the only place where the nonnegativity of an optional function is supposed. In practice only the following trivial special case is needed.

PROPOSITION 8.1. *If α is optional and t is a positive constant, then $\alpha + t$ is strictly optional.*

PROPOSITION 9. *If α is optional, β is arbitrary and $\alpha \leqq \beta$, then*

$$\mathcal{F}_{\alpha-} \subset \mathcal{F}_{\beta-}, \qquad \mathcal{F}_{\alpha+} \subset \mathcal{F}_{\beta+}. \tag{15}$$

Proof. $\forall t$: $\{\alpha > t\} \in \mathcal{F}_{t+}$, hence (8) with $\mathcal{F}_t$ replaced by $\mathcal{F}_{t+}$ proves the first relation in (15). Now $\forall \delta > 0$; $\alpha + \delta$ is (strictly) optional, hence

$$\mathcal{F}_{(\alpha+\delta)-} \subset \mathcal{F}_{(\beta+\delta)-}.$$

The second relation in (15) follows from this and Definition 2.

Definition 4. For any function α, let $\mathcal{F}_{\alpha(+)}$ $[\mathcal{F}_\alpha]$ denote the collection of subsets Λ of Ω for which $\Lambda \cap \{\alpha = +\infty\} \in \mathcal{F}_{+\infty}$ and (16) [(17)] is true:

$$\forall t \in T \colon \Lambda \cap \{\alpha < t\} \in \mathcal{F}_t; \tag{16}$$

$$\forall t \in T \colon \Lambda \cap \{\alpha \leqq t\} \in \mathcal{F}_t. \tag{17}$$

Clearly $\mathcal{F}_\alpha \subset \mathcal{F}_{\alpha(+)} \subset \mathcal{F}_{+\infty}$.

PROPOSITION 10. *If α is optional, then $\mathcal{F}_{\alpha(+)}$ is a B.F.; if α is strictly optional, then also is $\mathcal{F}_\alpha$. An equivalent definition of $\mathcal{F}_{\alpha(+)}$ is obtained if we replace $\mathcal{F}_t$ in (16) by $\mathcal{F}_{t-}$ or $\mathcal{F}_{t+}$; or if we replace $\mathcal{F}_t$ in (17) by $\mathcal{F}_{t+}$; or if we replace T by any dense subset of T.*

Proof. The first sentence follows from (10) and (11) by taking $\Lambda = \Omega$. The rest is similar to Proposition 1.

The following special case of the B.F.'s is instructive; the simple proof will be omitted.

PROPOSITION 11. *If α has a countable range $C \subset T$, then α is [strictly] optional if and only if*

$$\forall c \in C \colon \{\alpha = c\} \in \mathcal{F}_{c+}[\mathcal{F}_c];$$

and $\mathcal{F}_{\alpha(+)}[\mathcal{F}_\alpha, \mathcal{F}_{\alpha-}]$ is the collection of all sets Λ of the form:

$$\Lambda = \bigcup_{c \in C} \{\alpha = c\} \cap \Lambda_c$$

where $\Lambda_c \in \mathcal{F}_{c+}[\mathcal{F}_c, \mathcal{F}_{c-}]$.

PROPOSITION 12. *If* $\forall n$: α_n *is optional, and* $\alpha_n \downarrow \alpha$, *then*

$$\mathcal{F}_{\alpha(+)} = \bigwedge_n \mathcal{F}_{\alpha_n(+)}. \tag{18}$$

Remark. This is the counterpart of (7); note that (7) is valid if $\alpha_n \uparrow \alpha$ and each α_n is optional, by (5) and (15).

Proof. If α and β are optional, and $\alpha \leqq \beta$, $\Lambda \in \mathcal{F}_{\alpha(+)}$, then

$$\Lambda \cap \{\beta < t\} = [\Lambda \cap \{\alpha < t\}] \cap \{\beta < t\} \in \mathcal{F}_t,$$

by (16) and (10) for β. Hence $\mathcal{F}_{\alpha(+)} \subset \mathcal{F}_{\beta(+)}$. Applying this result to α and α_n, since α is optional by Proposition 7, we obtain " $\subset$ " instead of " $=$ " in (18). On the other hand, if Λ belongs to the right member of (18), then

$$\Lambda \cap \{\alpha < t\} = \bigcup_n [\Lambda \cap \{\alpha_n < t\}] \in \mathcal{F}_t.$$

Hence $\Lambda \in \mathcal{F}_{\alpha(+)}$ and (18) is proved.

PROPOSITION 13. *If* α *is arbitrary,* β *is* [*strictly*] *optional, and* $\alpha \leqq \beta$, *then*

$$\mathcal{F}_{\alpha-} \subset \mathcal{F}_{\beta(+)}[\mathcal{F}_\beta]. \tag{19}$$

If α *is optional,* β *is arbitrary, and* $\alpha < \beta$,[2] *then*

$$\mathcal{F}_{\alpha(+)} \subset \mathcal{F}_{\beta-}. \tag{20}$$

Proof. The first assertion is proved by the formula:

$$[\mathcal{F}_t \cap \{t < \alpha\}] \cap \{\beta < u\} = [\mathcal{F}_t \cap \{t < \alpha < u\}] \cap \{\beta < u\},$$

which is a collection of sets contained in $\mathcal{F}_u$, by (6) if $t < u$, and trivially if $t \geqq u$. Hence each generating set of $\mathcal{F}_{\alpha-}$ belongs to $\mathcal{F}_{\beta(+)}$ by definition, proving (19); similarly for the strict case. To prove (20), let $\Lambda \in \mathcal{F}_{\alpha+}$ and put

$$\Lambda_r = \Lambda \cap \{\alpha < r\}. \tag{21}$$

Then $\Lambda_r \in \mathcal{F}_r$ and consequently

$$\Lambda = \bigcup_{r \in R} [\Lambda \cap \{\alpha < r < \beta\}] = \bigcup_{r \in R} [\Lambda_r \cap \{r < \beta\}] \in \mathcal{F}_{\beta-}.$$

PROPOSITION 14. *If* α *is optional, then*

$$\mathcal{F}_{\alpha(+)} = \mathcal{F}_{\alpha+}.$$

---

[2] This means $\alpha < \beta$ on $\{\alpha < +\infty\}$ and $\alpha = \beta$ on $\{\alpha = +\infty\}$.

Proof. We have by Proposition 12 and the first part of the Proposition 13:

$$\mathcal{F}_{\alpha(+)} = \bigwedge_{n=1}^{\infty} \mathcal{F}_{(\alpha+n^{-1})(+)} \supset \bigwedge_{n=1}^{\infty} \mathcal{F}_{(\alpha+n^{-1})-}.$$

On the other hand, by the second part of Proposition 13:

$$\mathcal{F}_{\alpha(+)} \subset \bigwedge_{n=1}^{\infty} \mathcal{F}_{(\alpha+n^{-1})-}.$$

Together we obtain

$$\mathcal{F}_{\alpha(+)} = \bigwedge_{n=1}^{\infty} \mathcal{F}_{(\alpha+n^{-1})-} = \mathcal{F}_{\alpha+}.$$

Proposition 14 is useful since one or the other definition proves more convenient in application. From now on we shall drop the notation $\mathcal{F}_{\alpha(+)}$ in favor of $\mathcal{F}_{\alpha+}$ (which is defined for every α). Let us remark that for a strictly optional α, we have $\mathcal{F}_{\alpha*} \subset \mathcal{F}_{\alpha}$ but $\mathcal{F}_{\alpha}$ need not coincide with $\mathcal{F}_{\alpha*}$ (Example 2 in § 5); we have $\mathcal{F}_{\alpha} \subset \mathcal{F}_{\beta}$ if β is also strictly optional and $\alpha \leqq \beta$.

PROPOSITION 15. *If α is optional, $\beta \in \mathcal{F}_{\alpha+}$ and $\beta \geqq \alpha$, then β is optional. If furthermore either $\beta > \alpha$,*[3] *or α is strictly optional and $\beta \in \mathcal{F}_{\alpha}$, then β is strictly optional.*

Proof. We prove the first strict version. $\forall t$: $\{\beta \leqq t\} \in \mathcal{F}_{\alpha+}$, hence

$$\{\beta \leqq t\} = \{\beta \leqq t\} \cap \{\alpha < t\} \in \mathcal{F}_t$$

by Definition 4. Hence β is strictly optional.

PROPOSITION 15.1. *If α is [strictly] optional, $\Delta \in \mathcal{F}_{\alpha+}[\mathcal{F}_{\alpha}]$, and*

$$\alpha_\Delta = \begin{cases} \alpha & \text{on } \Delta, \\ +\infty & \text{on } \Omega - \Delta; \end{cases}$$

then α_Δ is [strictly] optional.

For an application of the preceding proposition, see e.g. [2; Lemma, p. 34].

An important special case of Proposition 15 is as follows. For an arbitrary optional α, we define

$$\alpha_n = \frac{[2^n\alpha + 1]}{2^n}, \qquad n \in N^0. \tag{22}$$

Let Z be the set $\{m\, 2^{-n}\}$ where n ranges over N^0 and m over N. Then $\forall n$: α_n

[3] See preceding footnote.

has a countable range contained in Z and is strictly optional. Hence by Proposition 11, $\mathcal{F}_{\alpha_n}$ consists of sets of the form

$$\bigcup_m [\{\alpha_n = m\,2^{-n}\} \cap \Lambda_{m2^{-n}}] = \bigcup_m [\{(m-1)2^{-n} \leqq \alpha < m\,2^{-n}\} \cap \Lambda_{m2^{-n}}]$$

where we use the notation (21). By the choice of the sequence $\{2^n\}$, we have $\alpha_{n+1} \leqq \alpha_n$ so that $\mathcal{F}_{\alpha_{n+1}} \subset \mathcal{F}_{\alpha_n}$, and $\mathcal{F}_{\alpha+}$ is the B. F. intersection of the nonincreasing sequence $\{\mathcal{F}_{\alpha_n}\}$, by (18). This type of approximation is useful in the evaluation of probabilities; see e. g. [1; pp. 165, 169].

2. Lattice and addition properties.

PROPOSITION 16. *If α is optional and β arbitrary, then*

$$\{\alpha < \beta\} \cap \mathcal{F}_{\alpha+} \subset \mathcal{F}_{\beta-}, \qquad \{\alpha \leqq \beta\} \cap \mathcal{F}_{\alpha+} \subset \mathcal{F}_{\beta+}. \tag{23}$$

Proof. Let $\Lambda \in \mathcal{F}_{\alpha+}$ and use the notation (21). We have

$$\Lambda \cap \{\alpha < \beta\} = \bigcup_{r \in R} [\Lambda \cap \{\alpha < r < \beta\}] = \bigcup_{r \in R} [\Lambda_r \cap \{r < \beta\}] \in \mathcal{F}_{\beta-}.$$

Next, $\forall m$:

$$\Lambda \cap \{\alpha \leqq \beta\} = \bigcap_{n=m}^{\infty} [\Lambda \cap \{\alpha < \beta + 1/n\}] \in \mathcal{F}_{(\beta+m^{-1})-};$$

hence $\Lambda \cap \{\alpha \leqq \beta\}$ belongs to $\mathcal{F}_{\beta+}$ by Definition 2.

PROPOSITION 17. *If α is optional and β arbitrary, then both $\{\alpha < \beta\}$ and $\{\alpha \leqq \beta\}$ belong to $\mathcal{F}_{(\alpha\wedge\beta)+}$.*

Proof. Since

$$\{\alpha \leqq \beta\} = \{\alpha \leqq (\alpha \wedge \beta)\},$$

this set belongs to $\mathcal{F}_{(\alpha\wedge\beta)+}$ by the second relation in (23). Applying this result to α and $\beta - n^{-1}$, we obtain

$$\{\alpha < \beta\} = \bigcup_n \{\alpha \leqq \beta - 1/n\} \in \bigvee_n \mathcal{F}_{(\alpha\wedge(\beta-n^{-1}))+} \subset \mathcal{F}_{(\alpha\wedge\beta)+}.$$

PROPOSITION 18. *If α and β are optional, then*

$$\{\alpha \leqq \beta\} \cap \mathcal{F}_{(\alpha\wedge\beta)\pm} = \{\alpha \leqq \beta\} \cap \mathcal{F}_{\alpha\pm}; \tag{24$\pm$}$$

$$\{\alpha \leqq \beta\} \cap \mathcal{F}_{(\alpha\vee\beta)\pm} = \{\alpha \leqq \beta\} \cap \mathcal{F}_{\beta\pm}, \tag{25$\pm$}$$

where in each formula we take "+" or "—" together, and similar relations also hold if "$\leqq$" is replaced by "$<$" everywhere.

Proof. It follows from (15) that we have " $\subset$ " in (24) and " $\supset$ " in (25). To prove " $\supset$ " in (24—), we need only observe that

$$\{\alpha \leqq \beta\} \cap [\mathcal{F}_t \cap \{t < \alpha\}] = \{\alpha \leqq \beta\} \cap [\mathcal{F}_t \cap \{t < (\alpha \wedge \beta)\}];$$

similarly for (25—), and with " $<$ " in place of " $\leqq$." To prove (24+) and (25+), we observe that, e. g.,

$$\{\alpha \leqq \beta\} \cap \mathcal{F}_{(\alpha \wedge \beta)+} = \{\alpha + \delta \leqq \beta + \delta\} \cap \bigwedge_{\delta > 0} \mathcal{F}_{[(\alpha \wedge \beta)+\delta]-}.$$

Applying (24—) and (25—) to $\alpha + \delta$ and $\beta + \delta$, we obtain (24+) and (25+) by the following general result, which is a counterpart to (4). If $\Delta \subset \Omega$, and $\{\mathcal{F}_i\}$ is a nonincreasing sequence of B. F.'s, then

$$\bigwedge_i (\Delta \mathcal{F}_i) = \Delta \cap (\bigwedge_i \mathcal{F}_i). \tag{26}$$

Proposition 18 is proved.

If $\mathcal{B}_1$ and $\mathcal{B}_2$ are two collections of subsets of Ω, we denote by $\mathcal{B}_1 \cup \mathcal{B}_2$ the collection of all sets of the form $\Lambda_1 \cup \Lambda_2$ where $\Lambda_1 \in \mathcal{B}_1$, $\Lambda_2 \in \mathcal{B}_2$.

PROPOSITION 19. *If α and β are optional, then*

$$\mathcal{F}_{(\alpha \vee \beta)+} = \mathcal{F}_{\alpha+} \cup \mathcal{F}_{\beta+}, \tag{27}$$

$$\mathcal{F}_{(\alpha \wedge \beta)+} = \mathcal{F}_{\alpha+} \wedge \mathcal{F}_{\beta+}, \tag{28}$$

$$\mathcal{F}_{(\alpha \vee \beta)-} = \mathcal{F}_{\alpha-} \vee \mathcal{F}_{\beta-}, \tag{29}$$

but in general

$$\mathcal{F}_{(\alpha \wedge \beta)-} \neq \mathcal{F}_{\alpha-} \wedge \mathcal{F}_{\beta-}. \tag{30}$$

Remark. These relations extend at once to a finite number of optional functions by induction. Previous relations (7) and (18) are the limiting cases of (29) and (28) respectively. The limiting case of (27) is in general false, see Example 7 in § 5.

Proof. It follows from (15) that we have " $\supset$ " in (27) and (29), " $\subset$ " in (28) and (30). It remains to prove the opposite inclusion in the first three relations and disprove it in the fourth.

Let $\Lambda \in \mathcal{F}_{(\alpha \vee \beta)+}$, then by (25+) and Proposition 16:

$$\{\alpha \leqq \beta\} \cap \Lambda \in \{\alpha \leqq \beta\} \cap \mathcal{F}_{\beta+} \subset \mathcal{F}_{\beta+}.$$

Interchanging α and β in the above and taking the union of the two results we obtain (27) with " $\subset$."

Next, let $\Lambda \in \mathcal{F}_{\alpha+}$, then by (24+) and Proposition 17:

$$\{\alpha \leqq \beta\} \cap \Lambda \in \{\alpha \leqq \beta\} \cap \mathcal{F}_{(\alpha \wedge \beta)+} \subset \mathcal{F}_{(\alpha \wedge \beta)+}.$$

If also $\Lambda \in \mathcal{F}_{\beta+}$, then we may interchange α and β in the above and take the union. Hence each Λ belonging to the right member of (28) belongs also to the left member.

Next, we observe that by (5):

$$\{\alpha \leqq \beta\} \in \mathcal{F}\{\alpha\} \vee \mathcal{F}\{\beta\} \subset \mathcal{F}_{\alpha-} \vee \mathcal{F}_{\beta-}.$$

Hence if $\Lambda \in \mathcal{F}_{(\alpha \vee \beta)-}$, then by (25—):

$$\{\alpha \leqq \beta\} \cap \Lambda \in \{\alpha \leqq \beta\} \cap \mathcal{F}_{\beta-} \subset \mathcal{F}_{\alpha-} \vee \mathcal{F}_{\beta-}.$$

Interchanging α and β we conclude as before that Λ belongs to the right member of (29).

Finally, an example of (30) will be given in Example 3 of § 5.

If α is optional, then so is $\alpha \wedge t$ for each $t \in T$, and $\alpha + t$ is strictly optional for each $t \in T^0$. Given $\{\mathcal{F}_t, t \in T\}$ and α, let us write:

(31) $$\mathcal{E}_t = \mathcal{F}_{(\alpha \wedge t)+}, \qquad \mathcal{G}_t = \mathcal{F}_{\alpha+t}.$$

The two families $\{\mathcal{E}_t, t \in T\}$ and $\{\mathcal{G}_t, t \in T^0\}$ are both nondecreasing.

PROPOSITION 20. *If α is optional, then*

(32±) $$\bigvee_t \mathcal{F}_{(\alpha \wedge t)\pm} = \mathcal{F}_{\alpha\pm}.$$

Proof. Since $\lim_{t \uparrow \infty} (\alpha \wedge t) = \alpha$, (32—) is just a special case of (7). However, the analogue of (7) with "—" replaced by "+" is in general false; see Example 7 in § 5.

To prove (32+), let $\Lambda \in \mathcal{F}_{\alpha+}$. Using the notation (21), we have for every $t \in T$:

$$\Lambda_n \cap \{(\alpha \wedge n) < t\} = \Lambda \cap \{\alpha < (n \wedge t)\} \in \mathcal{F}_{n \wedge t} \subset \mathcal{F}_t;$$

hence for every $n \in N^0$:

(33) $$\Lambda_n \in \mathcal{F}_{(\alpha \wedge n)+}.$$

Next, we have

(34) $$\{\alpha = +\infty\} \cap \mathcal{F}_{+\infty} \subset \mathcal{F}_{\alpha-}.$$

For if $M \in \mathcal{F}_n$ then $\{\alpha > n\} \cap M \in \mathcal{F}_{\alpha-}$ and so

(35) $$\{\alpha = +\infty\} \cap M = \bigcap_n [\{\alpha > n\} \cap M] \in \mathcal{F}_{\alpha-}.$$

Since $\alpha \in \mathcal{F}_{\alpha-}$, the class of sets M for which (35) holds forms a B. F. This B. F. contains every $\mathcal{F}_n$ as just shown, hence it contains $\mathcal{F}_{+\infty}$, proving (34). Since also $\Lambda \in \mathcal{F}_{+\infty}$, it follows that $\{\alpha = +\infty\} \cap \Lambda$ belongs to $\mathcal{F}_{\alpha-}$, and consequently to the left member of (32—), by (32—), and *a fortiori* to that of (32+). Now it follows from this, (33) and the equation

$$\Lambda = [\{\alpha = +\infty\} \cap \Lambda] \cup (\bigcup_{n=1}^{\infty} \Lambda_n),$$

that Λ belongs to the left member of (32+). Thus we have "$\supset$" in (32+). The opposite inclusion follows from (15) and so (32+) is proved.

PROPOSITION 21. *If β as well as α is optional relative to $\{\mathcal{F}_t\}$, then β is optional relative to $\{\mathcal{E}_t\}$ if and only if $\beta \in \mathcal{F}_{\alpha+}$. This is the case if $\beta \leqq \alpha$.*

Proof. If β is optional relative to $\{\mathcal{E}_t\}$, then by (28):

$$\beta \in \mathcal{E}_t = \mathcal{F}_{\alpha+} \wedge \mathcal{F}_{t+} \subset \mathcal{F}_{\alpha+}.$$

Conversely if $\beta \in \mathcal{F}_{\alpha+}$, then $\{\beta < t\} \in \mathcal{F}_{t+}$, and so

$$\{\beta < t\} \in \mathcal{F}_{\alpha+} \wedge \mathcal{F}_{t+} = \mathcal{F}_{(\alpha \wedge t)+}.$$

If $\beta \leqq \alpha$, then $\beta \in \mathcal{F}_{\beta+} \subset \mathcal{F}_{\alpha+}$ by (15). Proposition 21 is proved.

PROPOSITION 22. *If β is optional relative to $\{\mathcal{E}_t\}$, then*

$$\mathcal{E}_{\beta+} = \mathcal{F}_{(\alpha\wedge\beta)+}. \tag{36}$$

Proof. Let $\Lambda \in \mathcal{F}_{(\alpha\wedge\beta)+}$, then for every t:

$$\Lambda \cap \{\beta < t\} = [\Lambda \cap \{(\alpha \wedge \beta) < t\}] \cap \{\beta < t\} \in \mathcal{F}_t \cap \{\beta < t\} \subset \mathcal{F}_t.$$

By Proposition 21, $\beta \in \mathcal{F}_{\alpha+}$; also $\Lambda \in \mathcal{F}_{\alpha+}$ by (15), hence

$$\Lambda \cap \{\beta < t\} \in \mathcal{F}_{\alpha+}.$$

Combining the two relations above we obtain

$$\Lambda \cap \{\beta < t\} \in \mathcal{F}_{\alpha+} \wedge \mathcal{F}_{t+} = \mathcal{F}_{(\alpha\wedge t)+} = \mathcal{E}_t.$$

Hence $\Lambda \in \mathcal{E}_{\beta+}$ by definition.

Conversely, let $\Lambda \in \mathcal{E}_{\beta+}$, then by definition

$$\forall t: \quad \forall u: \quad \Lambda \cap \{\beta < t\} \cap \{(\alpha \wedge t) < u\} \in \mathcal{F}_u. \tag{37}$$

This reduces to the following two relations:

$$\forall t < u: \quad \Lambda \cap \{\beta < t\} \in \mathcal{F}_u; \tag{37'}$$

$$\forall t \geqq u: \quad \Lambda \cap \{\beta < t\} \cap \{\alpha < u\} \in \mathcal{F}_u. \tag{37''}$$

By (37′) and the optionality of α, we have

$$\forall t < u: \quad \Lambda \cap \{\beta < t\} \cap \{\alpha < u\} \in \mathcal{F}_u.$$

Combining this with (37″), and letting $t \uparrow \infty$, we obtain

$$(38) \qquad \forall u: \quad \Lambda \cap \{\beta < \infty\} \cap \{\alpha < u\} \in \mathcal{F}_u.$$

Since $\Lambda \in \mathcal{E}_{\beta+}$, we have also by (32+):

$$\Lambda \cap \{\beta = +\infty\} \in \mathcal{E}_{+\infty} = \mathcal{F}_{\alpha+}$$

and consequently

$$\forall u: \quad \Lambda \cap \{\beta = +\infty\} \cap \{\alpha < u\} \in \mathcal{F}_u.$$

Combining this with (38), we obtain

$$(39) \qquad \forall u: \quad \Lambda \cap \{\alpha < u\} \in \mathcal{F}_u.$$

Finally, since

$$(40) \qquad \Lambda \cap \{(\alpha \wedge \beta) < u\} = [\Lambda \cap \{\alpha < u\}] \cup [\Lambda \cap \{\beta < u\}],$$

and both members of the union above belong to $\mathcal{F}_u$ by (39) and (37′), we conclude that the left member of (40) belongs to $\mathcal{F}_u$. Thus $\Lambda \in \mathcal{F}_{(\alpha \wedge \beta)+}$ and Proposition 22 is proved. (In the final step we may also apply (28).)

The interest of Propositions 21 and 22 lies in this: given the optional β, any optional α dominating β can be made to play the role of $+\infty$ if the new family $\{\mathcal{E}_t\}$ is used instead of $\{\mathcal{F}_t\}$. In particular, considerations of (β, α) may be reduced to that of $(\beta, +\infty)$.

PROPOSITION 23. *If α is optional relative to $\{\mathcal{F}_t\}$ and $\beta \geqq \alpha$, then β is optional relative to $\{\mathcal{F}_t\}$ if and only if $\beta - \alpha$ is optional relative to $\{\mathcal{G}_t, t \in T^0\}$.*

Proof. If β is optional relative to $\{\mathcal{F}_t\}$, then we have by Proposition 16:

$$\forall t \in T^0: \quad \{\beta - \alpha < t\} = \{\beta < \alpha + t\} \in \mathcal{F}_{(\alpha+t)-} = \mathcal{G}_{t-}.$$

Hence $\beta - \alpha$ is optional relative to $\{\mathcal{G}_t\}$. Conversely if this is true, then

$$\forall s \in T^0: \quad \{\beta - \alpha < s\} \in \mathcal{G}_s = \mathcal{F}_{\alpha+s}.$$

Since $\{\alpha < r\} = \{\alpha + s < r + s\}$, and $\alpha + s$ is strictly optional, we have by Definition 4:

$$\forall r \in T, s \in T^0: \quad \{\beta - \alpha < s\} \cap \{\alpha < r\} \in \mathcal{F}_{r+s}.$$

It follows that

$$\forall t \in T: \quad \{\beta < t\} = \{\alpha + (\beta - \alpha) < t\} = \bigcup_{\substack{r+s<t \\ r \in R, s \in R^0}} \{\alpha < r; \beta - \alpha < s\} \in \mathcal{F}_t.$$

Hence β is optional relative to $\{\mathcal{F}_t\}$.

An application we may deduce, e.g. Lemma 4.4 in [4; p. 167].

PROPOSITION 23.1. *For any optional α and constant $t_0 \in T$, $(t_0 - \alpha) \vee 0$ is optional relative to $\{\mathcal{G}_t, t \in T^0\}$ and $(\alpha - t_0) \vee 0$ is optional relative to $\{\mathcal{F}_{t_0+t}, t \in T^0\}$.*

PROPOSITION 24. *If α and β are both optional relative to $\{\mathcal{F}_t\}$ and $\alpha \leqq \beta$, then*

$$\mathcal{G}_{(\beta-\alpha)+} = \mathcal{F}_{\beta+}.$$

Proof. Let $\Lambda \in \mathcal{F}_{\beta+}$, then we have by Proposition 16, for every $t \in T^0$:

$$\Lambda \cap \{\beta - \alpha < t\} = \Lambda \cap \{\beta < \alpha + t\} \in \mathcal{F}_{(\alpha+t)-} \subset \mathcal{F}_{\alpha+t} = \mathcal{G}_t.$$

Thus $\Lambda \in \mathcal{G}_{(\beta-\alpha)+}$ since $\beta - \alpha$ is optional relative to $\{\mathcal{G}_t\}$ by Proposition 23. Conversely, let $\Lambda \in \mathcal{G}_{(\beta-\alpha)+}$, then we have for every $t \in T$:

$$\Lambda \cap \{\beta < t\} = \bigcup_{\substack{r+s<t \\ r \in R,\ s \in R^0}} [\Lambda \cap \{\beta - \alpha < s\} \cap \{\alpha < r\}]$$

As in the preceding proof, the last-written union belongs to $\mathcal{F}_t$ and so $\Lambda \in \mathcal{F}_{\beta+}$.

3. Progressive and natural Borel measurability. Let X be a space, $\mathcal{A}$ a Borel field on X. For each $t \in T$, let $x_t: \omega \to x_t(\omega)$ be a function on Ω to X. We write also $x(t,\omega)$ for $x_t(\omega)$. For each t, it is clear that $x_t^{-1}(\mathcal{A})$ is a B.F. on Ω. Let

$$\mathcal{F}_t^0 = \bigvee_{s \leqq t} x_s^{-1}(\mathcal{A}).$$

Then $\mathcal{F}_t^0$ is the B.F. generated by all x_s with $s \leqq t$, and $\{\mathcal{F}_t^0, t \in T\}$ is a nondecreasing family of B.F.'s on Ω.

Given $\{x_t\}$, the family of B.F.'s $\{\mathcal{F}_t\}$ on Ω is said to be *adapted to* $\{x_t\}$ iff it is nondecreasing and $x_t \in \mathcal{F}_t$ for each t. The family $\{\mathcal{F}_t^0\}$ defined above is adapted and is minimal in the sense that for any adapted family $\{\mathcal{F}_t\}$ we have $\mathcal{F}_t^0 \subset \mathcal{F}_t$ for each t.

Definition 5. The family $\{\mathcal{F}_t^0\}$ is called the *natural family* of Borel fields for $\{x_t\}$.

If α is a function on Ω to T, then the function $\omega \to x(\alpha(\omega), \omega)$ will sometimes be denoted by x_α. In particular we shall write:

$$x_{\alpha \wedge t}: \quad \omega \to x(\alpha(\omega) \wedge t, \omega),$$
$$x_{\alpha+t}: \quad \omega \to x(\alpha(\omega) + t, \omega).$$

Note that α is supposed finite here since $x_{\pm\infty}$ have not been defined.

Let θ be an element alien to $X: \theta \notin X$. Put $X_0 = X \cup \{\theta\}$ and define $\mathcal{A}_\theta$ to be the Borel field generated by $\mathcal{A}$ and the singleton $\{\theta\}$. Now for given $\{x_t\}$ and α define:

$$\forall t \in T: \quad x_t^- = \begin{cases} x_t & \text{on } \{t < \alpha\}, \\ \theta & \text{on } \{t \geqq \alpha\}. \end{cases} \tag{42}$$

Let $\mathcal{F}^-$ be the B. F. generated by $\{x_t^-, t \in T\}$. Recall the definition of $\mathcal{F}_{\alpha-}^0$ from Definitions 1 and 5.

PROPOSITION 25. *We have* $\mathcal{F}^- = \mathcal{F}_{\alpha-}^0$.

Proof. For each t, we have

$$\{x_t^- = \theta\} = \{\alpha \leqq t\}, \tag{43}$$

hence $\alpha \in \mathcal{F}^-$. For any A in $\mathcal{A}$ and $s \leqq t$, we have

$$\{x_s \in A\,; t < \alpha\} = \{x_s^- \in A\}, \tag{44}$$

from which we deduce the more general relation

$$\mathcal{F}_t^0 \cap \{t < \alpha\} \subset \mathcal{F}^-.$$

Hence each generating set of $\mathcal{F}_{\alpha-}^0$ belongs to $\mathcal{F}^-$ and so $\mathcal{F}_{\alpha-}^0 \subset \mathcal{F}^-$. Conversely, we see from (43) that

$$\{x_t^- = \theta\} \in \mathcal{F}\{\alpha\} \subset \mathcal{F}_{\alpha-}^0,$$

by (5), and from (44) that

$$\{x_t^- \in A\} \in \mathcal{F}_{\alpha-}^0.$$

It follows from the definition of $\mathcal{A}_\theta$ that $x_t^- \in \mathcal{F}_{\alpha-}^0$ so that $\mathcal{F}^- \subset \mathcal{F}_{\alpha-}^0$. Proposition 25 is proved.

Remark. If we replace $\{t < \alpha\}$ and $\{t \geqq \alpha\}$ in (42) respectively by $\{t \leqq \alpha\}$ and $\{t > \alpha\}$, call the resulting function x_t^*, and $\mathcal{F}^*$ the B. F. generated by $\{x_t^*, t \in T\}$, then we have

$$\mathcal{F}_{\alpha-}^0 = \mathcal{F}^- \subset \mathcal{F}^* = \mathcal{F}_{\alpha\cdot}^0$$

where the last B. F. is defined in the Remark after Definition 2, and the "$\subset$" above cannot be replaced by "$=$" in general (Example 6 in § 5). A similar B. F., that generated by $\{x_{\alpha\wedge t}, t \in T\}$, has been used in very special cases such as Brownian motion to play the role of $\mathcal{F}_{\alpha-}^0$. However, this new field may not contain α (Example 4 in § 5) but must contain x_α, which need not be contained in $\mathcal{F}_{\alpha-}$, nor even in $\mathcal{F}_{\alpha+}$ for optional α (Example

5 in § 5). Indeed, we shall proceed to find a condition under which $x_\alpha \in \mathcal{F}_{\alpha+}$ for every optional α.

As a matter of general terminology, if X_i is a set and $\mathcal{F}_{X_i}$ is a B. F. on X_i for $i = 1, 2$, we shall write $f \in \mathcal{F}_{X_1}/\mathcal{F}_{X_2}$ iff $f^{-1}(\mathcal{F}_{X_2}) \subset \mathcal{F}_{X_1}$. (Thus our earlier notation e. g., "$\alpha \in \mathcal{F}_{\alpha-}$" is an abbreviation for "$\alpha \in \mathcal{F}_{\alpha-}/\mathcal{B}$.") The product space $X_1 \times X_2$ and product field $\mathcal{F}_{X_1} \times \mathcal{F}_{X_2}$ are defined as usual (see [5; pp. 150 ff.]). In this connection let us record a well-known result.

PROPOSITION 26. *Let* $(X_i, \mathcal{F}_{X_i})$, $i = 1, 2, 3$, *be three pairs of space-fields. Suppose that*

$$f \in \mathcal{F}_{X_1} \times \mathcal{F}_{X_2}/\mathcal{F}_{X_3}, \quad \phi_1 \in \mathcal{F}_{X_1} \times \mathcal{F}_{X_2}/\mathcal{F}_{X_1}, \quad \phi_2 \in \mathcal{F}_{X_1} \times \mathcal{F}_{X_2}/\mathcal{F}_{X_2},$$

and g *is the function*

$$g: (\xi_1, \xi_2) \to f(\phi_1(\xi_1, \xi_2), \phi_2(\xi_1, \xi_2)).$$

Then $g \in \mathcal{F}_{X_1} \times \mathcal{F}_{X_2}/\mathcal{F}_{X_3}$.

Now consider the pairs

$$(T, \mathcal{B}), \qquad (\Omega, \mathcal{F}), \qquad (X, \mathcal{A})$$

where $\mathcal{F}$ is an arbitrary B. F. on Ω and the other symbols have been introduced before. Suppose that for each $t \in T$, we have $x_t \in \mathcal{F}/\mathcal{A}$, then $\mathcal{F}_t^0 \subset \mathcal{F}$ and so $\mathcal{F}_{+\infty}^0 \subset \mathcal{F}$. In the usual language $\{x_t, t \in T\}$ is a family of measurable functions on the measurable space $(\Omega, \mathcal{F})$. The family $\{x_t\}$ is called *Borel measurable* iff the function $x(\cdot, \cdot) \in \mathcal{B} \times \mathcal{F}/\mathcal{A}$. The following definition is more stringent, and the rest of the section is devoted to developing its main consequences. For the applicability of the new concept see the following section.

Definition 6. The family $\{x_t\}$ is said to be *progressively Borel measurable relative to the adapted family* $\{\mathcal{F}_t\}$ iff we have

$$\forall a \in T, A \in \mathcal{A}: \{(t, \omega): t < a, x(t, \omega) \in A\} \in \mathcal{B}_a \times \mathcal{F}_a; \tag{45}$$

it is said to be *naturally Borel measurable* iff it is progressively Borel measurable relative to its natural family.

In symbols, (45) may be written as

$$x\,|_{T_{a-} \times \Omega} \in \mathcal{B}_a \times \mathcal{F}_a/\mathcal{A} \tag{45'}$$

where $T_{a-} = (-\infty, a)$ and $x\,|_{T_{a-} \times \Omega}$ is the restriction of $x(\cdot, \cdot)$ to $T_{a-} \times \Omega$. Let us recall that $T_a = (-\infty, a]$, $\mathcal{B}_a = T_a \cap \mathcal{B}$.

It follows from the next proposition that progressive measurability relative to $\{\mathcal{F}_t\}$ or $\{\mathcal{F}_{t-}\}$ or $\{\mathcal{F}_{t+}\}$ is the same concept.

PROPOSITION 27. *An equivalent definition of progressive measurability is obtained if we replace $\mathcal{F}_a$ in (45) by $\mathcal{F}_{a-}$ or $\mathcal{F}_{a+}$; or if we replace the right member of (45) by*

$$\bigwedge_{\delta>0}[\mathcal{B}_{a+\delta}\times\mathcal{F}_{a+\delta}]=\bigwedge_{\delta>0}[\mathcal{B}_a\times\mathcal{F}_{a+\delta}] \tag{46}$$

or if we do this and also replace "$t<a$" in (45) by "$t\leqq a$."

The proof is similar to that of Proposition 1 or 5. The equation (46) is not quite trivial but its proof will be omitted. It is not known if in general the B. F. in (46) coincides with $\mathcal{B}_a\times\mathcal{F}_{a+}$.[4]

For an arbitrary index set the product space X^I and product field $\mathcal{A}^I$ are defined in the usual way ([5; p. 158]); and the function ϕ on X^I to X will be called Borel measurable iff $\phi\in\mathcal{A}^I/\mathcal{A}$. (For certain topological B. F.'s $\mathcal{A}$ such a function is called a Baire function.)

PROPOSITION 28. *Let $\{x_t^{(i)}, t\in T\}$, $i\in I$, be a collection of families, progressively Borel measurable relative to the same adapted $\{\mathcal{F}_t, t\in T\}$ and let ϕ be a Borel measurable function on X^I to X:*

$$\phi:(\xi^{(i)}, i\in I)\to\phi(\xi^{(i)}, i\in I).$$

Then the family $\{\Phi_t, t\in T\}$, where Φ is the function on $\mathcal{B}\times\mathcal{F}$ to $\mathcal{A}$:

$$\Phi:(t,\omega)\to\phi(x^{(i)}(t,\omega), i\in I),$$

is progressively Borel measurable relative to $\{\mathcal{F}_t\}$.

This proposition, like Proposition 26, is an easy analogue of the classical result to the effect that "a Borel measurable function of Borel measurable functions is Borel measurable."

PROPOSITION 29. *Let $\{x_t\}$ be progressively Borel measurable relative to $\{\mathcal{F}_t\}$, and ϕ be a function on Ω to T_t such that $\phi\in\mathcal{F}_{t+}/\mathcal{B}_t$. If x_ϕ denotes the function*

$$x_\phi:\omega\to x(\phi(\omega),\omega),$$

then $x_\phi\in\mathcal{F}_{t+}/\mathcal{A}$.

Proof. Without using one of the equivalent definitions in Proposition 27, let us first suppose that ϕ is on Ω to T_{t-}. Consider the three pairs

$$(T_{t-},\mathcal{B}_t),\qquad(\Omega,\mathcal{F}_{t+}),\qquad(X,\mathcal{A}),$$

[4] However, P. A. Meyer has proved a result which implies the truth of this if the product fields are augmented by all null sets of a produce measure on $\mathcal{B}\times\mathcal{F}$.

and the three functions: $f \equiv x\,|_{T_t \times \Omega}$, $\phi_1(t,\omega) \equiv \phi(\omega)$, $\phi_2(t,\omega) \equiv \omega$. By (45′), $f \in \mathcal{B}_t \times \mathcal{F}_{t+}/\mathcal{A}$. Hence an application of Proposition 26 yields $x_\phi \in \mathcal{B}_t \times \mathcal{F}_{t+}/\mathcal{A}$. Since x_ϕ is a function of ω alone this reduces to $x_\phi \in \mathcal{F}_{t+}/\mathcal{A}$. Now if ϕ is on Ω to T_t, we replace t by $t+\delta$ in the above for $\delta > 0$ and we have by what has just been proved: $x_\phi \in \mathcal{F}_{t+\delta}/\mathcal{A}$. This being true for every $\delta > 0$, we obtain $x_\phi \in \mathcal{F}_{t+}/\mathcal{A}$ as asserted.

PROPOSITION 30. *Let $\{x_t\}$ be progressively Borel measurable relative to $\{\mathcal{F}_t\}$, and α be finite-valued and optional relative to $\{\mathcal{F}_t\}$. Then we have*

$$\forall t \in T: \quad x_{\alpha \wedge t} \in \mathcal{F}_{(\alpha \wedge t)+} \tag{47}$$

where the "$+$" may be omitted if α is strictly optional; and

$$x_\alpha \in \mathcal{F}_{\alpha+}. \tag{48}$$

Proof. Since $\alpha \wedge t$ is optional, we have $\alpha \wedge t \in \mathcal{F}_{(\alpha \wedge t)+}$; since also $\alpha \wedge t \leqq t$, we may apply Proposition 29 to obtain

$$\forall t \in T: \quad x_{\alpha \wedge t} \in \mathcal{F}_{t+}.$$

It follows that for any $A \in \mathcal{A}$ and $s \in T$, we have

$$\begin{aligned} &\{x_{\alpha \wedge t} \in A\} \cap \{\alpha < s\} \\ &\qquad = \{x_{\alpha \wedge (s \wedge t)} \in A\} \cap \{\alpha < s\} \in \mathcal{F}_{(s \wedge t)+} \vee \mathcal{F}_{s+} = \mathcal{F}_{s+}. \end{aligned}$$

Thus $x_{\alpha \wedge t} \in \mathcal{F}_{\alpha+}$ by Definition 4, and consequently by (28):

$$x_{\alpha \wedge t} \in \mathcal{F}_{\alpha+} \wedge \mathcal{F}_{t+} = \mathcal{F}_{(\alpha \wedge t)+}.$$

Hence (47) is proved and the case of strict optionality is similar. Furthermore, we have by (47) and (32+) (note that since α is finite here, (32+) follows simply from (33)):

$$x_\alpha = \lim_{t \uparrow \infty} x_{\alpha \wedge t} \in \bigvee_t \mathcal{F}_{(\alpha \wedge t)+} = \mathcal{F}_{\alpha+}.$$

Proposition 30 is proved.

Let $\alpha(\cdot,\cdot)$ be a function on $T \times \Omega$ to T with the following properties:

(i) $\forall t \in T$: $\alpha_t(\cdot) = \alpha(t,\cdot)$ is optional relative to $\{\mathcal{F}_s, s \in T\}$;

(ii) $\forall \omega \in \Omega$: $\alpha(\cdot,\omega)$ is nondecreasing and right continuous on T.

The family $\{\mathcal{F}_{\alpha_t+}, t \in T\}$ is then nondecreasing by (15), and adapted to $\{x_{\alpha_t}, t \in T\}$ by (48), where x_{α_t} is the function below:

$$x_{\alpha_t}: \omega \rightarrow x(\alpha(t,\omega),\omega).$$

Examples are, for an optional $\alpha(\cdot)$:

$$\alpha(t,\omega) \equiv \alpha(\omega) \wedge t, \qquad \alpha(\omega) \vee t, \qquad \alpha(\omega) + t,$$

the last for $t \in [0,\infty)$.

PROPOSITION 31. *If $\{x_t\}$ is progressively Borel measurable relative to $\{\mathcal{F}_t\}$, then so is $\{x_{\alpha_t}\}$ relative to $\{\mathcal{F}_{\alpha_t+}\}$.*

Proof. We prove first the following lemma which is the particular case of the proposition for $x(t,\omega) \equiv t$.

LEMMA. *The family $\{\alpha_t, t \in T\}$ (of functions on Ω to T, indexed by T) is progressively Borel measurable relative to $\{\mathcal{F}_{\alpha_t+}\}$.*

Proof of the Lemma. Let $a \in T$, $c \in T$, then it follows from the hypotheses in (ii) that

$$(49) \qquad \{(t,\omega): t < a, \alpha(t,\omega) < c\} = \bigcup_{r \in R \cap (-\infty,a)} \{(t,\omega): t < r, \alpha(r,\omega) < c\}.$$

Since $r < a$, we have $\alpha_r \leqq \alpha_a$ by (ii), and so the set in { } on the right side of (49) may be written as $T_{r-} \times F$ where

$$F = \{\omega: [\alpha_r(\omega) \wedge \alpha_a(\omega)] < c\}.$$

Since $F \in \mathcal{F}_{\alpha_a+}$ by (5) and (15), we see that the set on the left side of (49) belongs to $\mathcal{B}_a \times \mathcal{F}_{\alpha_a+}$. Hence $\{\alpha_t\}$ is progressively Borel measurable relative to $\{\mathcal{F}_{\alpha_t+}\}$ by definition. The Lemma is proved.

Next, we prove that for each $\Gamma \in \mathcal{B} \times \mathcal{F}_{\alpha_a+}$, we have

$$(50) \qquad \{(t,\omega): t < a, (\alpha(t,\omega),\omega) \in \Gamma\} \in \mathcal{B}_a \times \mathcal{F}_{\alpha_a+}.$$

It is sufficient to prove (50) for Γ of the form $B \times F$ where $B \in \mathcal{B}$ $F \in \mathcal{F}_{\alpha_a+}$. For such a set the left member of (50) reduces to

$$(51) \qquad (T_{a-} \times F) \cap \{(t,\omega): t < a; \alpha(t,\omega) \in B\}.$$

By the lemma above, the set in { } in (51) belongs to $\mathcal{B}_a \times \mathcal{F}_{\alpha_a+}$; since $T_{a-} \times F$ also belongs to this field, (50) follows.

Now let $A \in \mathcal{A}$, and define two subsets Γ_1 and Γ_2 of $T \times \Omega$ as follows:

$$(52) \qquad \Gamma_1 = \{(s,\omega): s < \alpha(a,\omega); x(s,\omega) \in A\},$$

$$(53) \qquad \Gamma_2 = \{(s,\omega): s = \alpha(a,\omega); x(s,\omega) \in A\}.$$

We have

$$(54) \qquad \Gamma_1 = \bigcup_{r \in R} \{(s,\omega): s < r < \alpha(a,\omega); x(s,\omega) \in A\}.$$

For each r the set on the right side of (54) may be written as

$$\{(s,\omega): s < r; x(s,\omega) \in A\} \cap (T \times M) \tag{55}$$

where

$$M = \{\omega: r < \alpha(a,\omega)\}.$$

By the progressive Borel measurability of $\{x_t\}$, the set in { } in (55) belongs to $\mathcal{B}_r \times \mathcal{F}_r$, and we are going to prove that

$$(\mathcal{B}_r \times \mathcal{F}_r) \cap (T \times M) \in \mathcal{B} \times \mathcal{F}_{\alpha_a+}. \tag{56}$$

To do this let $B \in \mathcal{B}_r$ and $F \in \mathcal{F}_r$, then

$$(B \times F) \cap (T \times M) = (B \cap T) \times (F \cap M). \tag{57}$$

Since

$$F \cap M \in \mathcal{F}_r \cap \{r < \alpha_a\} \subset \mathcal{F}_{\alpha_a-} \subset \mathcal{F}_{\alpha_a+}$$

by Definitions 1 and 2, the set in (57) belongs to the right member of (56). This is sufficient for the truth of (56) in general. Consequently every member of the union on the right side of (54) belong to $\mathcal{B} \times \mathcal{F}_{\alpha_a+}$, and so does Γ_1 also. As for Γ_2, we have

$$\Gamma_2 = \{(s,\omega): s = \alpha(a,\omega); x(\alpha(a,\omega),\omega) \in A\};$$

since $\alpha_a \in \mathcal{F}_{\alpha_a+}$ and $x_{\alpha_a} \in \mathcal{F}_{\alpha_a+}$ by (48), it is clear that Γ_2 belongs to $\mathcal{B} \times \mathcal{F}_{\alpha_a+}$. Hence so does $\Gamma_0 = \Gamma_1 \cup \Gamma_2$, and we may substitute Γ_0 for Γ in the left member of (50). The resulting relation, since $t < a$ implies $\alpha(t,\omega) \leqq \alpha(a,\omega)$, is

$$\{(t,\omega): t < a; x(\alpha(t,\omega),\omega) \in A\} \in \mathcal{B}_a \times \mathcal{F}_{\alpha_a+}$$

by (52) and (53). Since a and A are arbitrary, we have proved that $\{x_{\alpha_t}\}$ is progressively Borel measurable relative to $\{\mathcal{F}_{\alpha_t+}\}$.

4. **The introduction of measure.** In this section we shall show that the assumption of natural Borel measurability as defined in § 3 is a reasonable one for stochastic processes.

Consider the family $\{x_t, t \in T\}$ of measurable functions in the measurable space $(\Omega, \mathcal{F})$. If a probability measure P is given on $\mathcal{F}$, the family is called a *stochastic process* on the *probability space* $(\Omega, \mathcal{F}, P)$. A second process $\{x'_t, t \in T\}$ is called a *standard modification* of the first iff for every t:

$$P\{\omega: x_t(\omega) = x'_t(\omega)\} = 1.$$

(See [3, Ch. 2]).

We shall deal with functions from measurable spaces to metric spaces. The Borel sets of a metric space will be taken to be the sets in the Borel field generated by the closed sets. A function f from a measurable space to a metric space will be called Borel iff the inverse image under f of a Borel set is measurable. In particular f will be called simple if f assumes only finitely many values, each on a measurable set. The function f will be called *separably Borel measurable* iff it is in the smallest class of functions containing the simple functions and closed under pointwise convergence of sequences. A separably Borel measurable function is Borel measurable, and has a separable range. Conversely if a Borel measurable function has a separable range it is separably Borel measurable.

Let $(T, \mathcal{B})$, $(\Omega, \mathcal{F})$ be measurable spaces. In this paragraph T, $\mathcal{B}$ need not have the specific interpretation made elsewhere in the paper. Then $(T \times \Omega, \mathcal{B} \times \mathcal{F})$ is a measurable space. Let x be a function from this space into a metric space. Then x will be called a simple product-space function iff it is simple and if each value is assumed on a set which is a finite union of direct products of $\mathcal{B}$ and $\mathcal{F}$ sets. It is easily seen that the separably Borel measurable functions from $T \times \Omega$ into the metric space are those in the smallest class of functions containing the simple product-space functions and closed under pointwise convergence of sequences.

Throughout the rest of this section all processes will be "metric space valued" processes, by which we mean that X is a metric space (distance function ρ) and that $\mathcal{A}$ is the class of Borel subsets of X. The functions of the process are supposed not merely to belong to $\mathcal{F}/\mathcal{A}$ however, that is to be Borel measurable, but even to be separably Borel measurable. The process $\{x_t\}$ will be said to be *separably Borel measurable* iff $x(\cdot, \cdot)$ is separably Borel measurable; similarly when the adverb "progressively" or "naturally" is added.

Let M be the space of separably Borel measurable functions from Ω into X. If x and y are in M define the distance between them as

$$\inf_{\epsilon>0} [\epsilon + P\{\omega : \rho(x(\omega), y(\omega)) > \epsilon\}]$$

and identify x with y if they are equal almost everywhere. Then M becomes a metric space $\hat{M}$, and we shall denote by $\hat{x}$ the element of $\hat{M}$ corresponding to the function x in M. Convergence in the metric sense in $\hat{M}$ corresponds to convergence in measure in M. Moreover if the distance between $\hat{x}_n$ and $\hat{x}$ is the n-th term of a convergent series, $\lim_{n\to\infty} x_n = x$ almost everywhere. The process $\{x_t\}$ defines a function $\hat{x}_.$ from T into $\hat{M}$, and the properties of the process are reflected in those of the function $\hat{x}$.

PROPOSITION 32. *If the metric space valued process $\{x_t\}$ is separably Borel measurable, $\hat{x}_.$ is separably Borel measurable. Conversely if the function $\hat{x}_.$ is separably Borel measurable there is a standard modification of the process which is natuarlly separably Borel measurable.*

Proof. (The method of proof was suggested by P. A. Meyer.) Consider the class of processes $\{x_t\}$ corresponding to separably Borel measurable functions $x(\cdot,\cdot)$ for which the corresponding function $\hat{x}_.$ is separably Borel measurable. The class Γ of functions $x(\cdot,\cdot)$ so defined contains the simple product-space functions, because for such a function $\hat{x}_.$ becomes simple. The class Γ is closed under sequential convergence and therefore contains all separably Borel measurable functions. Thus the direct half of the proposition is true. Conversely suppose that the function $\hat{x}_.$, which we shall denote by $\hat{x}(\cdot)$, is separably Borel measurable. Then the range of this function is separable. Let $S_1^n, S_2^n, \cdots$ be a disjoint partition of the closure of this range into Borel sets of diameter $< 1/n^2$, with the $(n+1)$-th partition a refinement of the n-th. Define A_{ji}^n by

$$A_{ji}^n = \{t : \hat{x}(t) \in S_j^n, i2^{-n} < t \leqq (i+1)2^{-n}\}.$$

Then for each n, $\{A_{ji}^n\}$ is a partition of T. If A_{ji}^n is not empty choose a point t_{ji}^n in it, making the choice in such a way that each t_{ji}^n is also some t_{lk}^{n+1}. Define the function ϕ_n from T to T by

$$\phi_n(t) = t_{ji}^n \text{ on } A_{ji}^n.$$

Then ϕ_n is Borel measurable and

$$x[\phi_n(t)] = x(t) \qquad \text{if } t = t_{ji}^m \text{ and } n \geqq m.$$
$$|\phi_n(t) - t| < 2^{-n}.$$

Moreover for each value of t,

$$\lim_{n\to\infty} x[\phi_n(t),\omega] = x(t,\omega) \tag{58}$$

for almost all ω. Define $x_0(t,\omega)$ as the limit on the left when the limit exists and as c otherwise, where c is some specified element of X. Then $x_0(t,\omega) = x(t,\omega)$ for $t = t_{ji}^n$ and the $x_0(t)$ process is a standard modification of the given one, determined completely by $x(t,\omega)$ for $t \in \{t_{ji}^n\}$ and by the choice of c. If $\delta > 0$ and if n is sufficiently large, $x[\phi_n(t),\omega]$ defines a process whose restriction to the interval $(-\infty, a)$ is separably Borel measurable relative to the field $\mathcal{B}_{a+\delta} \times \mathcal{F}_{a+\delta}$ where $\{\mathcal{F}_t\}$ is the natural field family for the $x_0(t)$ process. Hence the same restriction of the $x_0(t,\omega)$ process

has the same measurability property for all $\delta > 0$. The $x_0(t,\omega)$ process is therefore progressively separably Borel measurable relative to its natural field family, as was to be shown.

We have actually proved more than the proposition asserts. The new natural fields are contained in the old ones for $t \in t_{j}{}^{n}$, a set dense in T. Moreover the assertions about the new process are also valid for its restriction to any interval of the form (b, ∞).

If we make the additional assumption that the range space X of the random variables is compact as well as metric, the conclusion of Proposition 32 can be strengthened. Let f be a function on T into X. For an arbitrary subset A of T, let $f[A]$ be the range of the restriction of f to A, and let $f[A]^*$ be the closure of $f[A]$ in the topology of X. For each t in T, we set

$$\begin{aligned} L_+(f,A,t) &= \bigcap_n f[[t,t+n^{-1}] \cap A]^*, \\ L_-(f,A,t) &= \bigcap_n f[[t-n^{-1},t] \cap A]^*, \\ L(f,A,t) &= \bigcap_n f[[t-n^{-1},t+n^{-1}] \cap A]^* \\ &= L_-(f,A,t) \cup L_+(f,A,t). \end{aligned}$$

Thus $L[L_+,L_-]$ is the set of [right, left] limiting values of f on A at t. The function f is said to be *[right, left] separable at t with A as a separating set* iff

$$f(t) \in L(f,A,t)\,[L_+(f,A,t), L_-(f,A,t)].$$

It is said to be [right, left] separable iff this is so at each t in T. The process $\{x_t, t \in T\}$ taking values in X is said to be *[right, left] separable* iff there is a countable set A such that for each ω in Ω, the sample function $x(\cdot,\omega)$ is so separable with A as a separating set.[5]

PROPOSITION 33. *If X is compact metric, the standard modification described in Proposition 32 can be made separable in addition to the other stated properties.*

Proof. To prove this assertion we change the definition of $x_0(t,\omega)$ in the proof of Proposition 32. Let $\{\xi_n\}$ be a sequence dense in X and define: $v_{11}(t,\omega)$ is the smallest value of $n \geqq 1$ for which

$$\rho(\xi_1, x[\phi_n(t),\omega]) \leqq \liminf_{m\to\infty} \rho(\xi_1, x[\phi_m(t),\omega]) + 1;$$

[5] These definitions were given by Chung in unpublished lecture notes in 1962, in which he proved that every real-valued process has a standard modification which is right [left] separable.

$v_{1j}(t,\omega)$ is the smallest value of $n > v_{1,j-1}(t,\omega)$ for which

$$\rho(\xi_1, x[\phi_n(t),\omega]) \leqq \liminf_{m\to\infty} \rho(\xi_1, x[\phi_m(t),\omega]) + 1/j.$$

Then

$$\lim_{j\to\infty} \rho(\xi_1, x[\phi_{v_{1j}}(t),\omega]) = \liminf_{m\to\infty} \rho(\xi_1, x[\phi_m(t),\omega])$$

and it is clear that if $\delta > 0$ the restriction of v_{1j} to $T_{a-}\times\Omega$ belongs to $\mathscr{B}_{a+\delta}\times\mathscr{F}_{a+\delta}$ for sufficiently large j. In general choose $\{v_{kj}(t,\omega), j \geqq 1\}$ a subsequence of $\{v_{k-1,j}(t,\omega), j \geqq 1\}$ in such a way that the restriction of v_{kj} to $T_{a-}\times\Omega$ belongs to $\mathscr{B}_{a+\delta}\times\mathscr{F}_{a+\delta}$ for sufficiently large j, whenever $\delta > 0$, and that

$$\lim_{j\to\infty} \rho(\xi_k, x[\phi_{v_{kj}}(t),\omega]) = \liminf_{j\to\infty} \rho(\xi_k, x[\phi_{v_{k-1,j}}(t),\omega])$$

Then

$$\lim_{j\to\infty} \rho(\xi_k, x[\phi_{v_{jj}}(t),\omega])$$

exists for all k, t, ω. Hence $\lim_{j\to\infty} x[\phi_{v_{jj}}(t),\omega]$ exist for all (t,ω) and we define $x_0(t)$ as this limit. Clearly the x_{0t} process is a standard modification of the x_t process with all the properties required in the proposition. The separating sequence is the sequence $\{t_{ji}{}^n\}$ defined in the proof of Proposition 32. Finally we note that the x_{0t} process is 'right separable' as defined above, if $\phi_n(t) \geqq t$. This inequality is not necessarily satisfied as we have defined ϕ_n, but the definition can be modified to achieve this inequality as follows. Define $A_{ji}{}^n$ as in the proof of Proposition 32. If $A_{ji}{}^n$ contains its supremum let $t_{ji}{}^n$ be this supremum and define $\phi_n(t) = t_{ji}{}^n$ on $A_{ji}{}^n$ as before. Otherwise let $\{t_{jik}{}^n, k \geqq 1\}$ be a monotone sequence in $A_{ji}{}^n$ with limit equal to this supremum and define

$$\phi_n(t) = t_{jik}{}^n \text{ on } A_{ji}{}^n \cap (t_{ji,k-1}{}^n, t_{jik}{}^n], \quad k \geqq 1 \text{ where } t_{ji0}{}^n = i2^{-n}.$$

As a complement we shall consider functions from a measure space to a metric space. The measure, say ν will be supposed complete. A function from the measure space to a metric space will be called ν-measurable iff it is the ν-almost everywhere limit of a sequence of simple functions. The range of the function is then ν-almost separable, that is, the restriction of the function to the complement of some set of ν-measure 0 is separable. If a function is separably Borel measurable it is ν-measurable and conversely a ν-measurable function coincides ν-almost everywhere with some separably Borel measurable function. A Borel measurable function is ν-measurable if and only if it is ν-almost separably valued. The same assertions are true if Borel

measurability is defined not using the domain of ν but any Borel subfield whose completion under ν yields the given domain of ν. In the following, μ is any Lebesgue-Stieltjes measure on $(-\infty, +\infty)$, that is, a completed measure of Borel sets, finite for compact sets. Let $\mathcal{B}^*$ be the domain of μ. The completed product measure $\mu \times P$ is defined on an extension of $\mathcal{B}^* \times \mathcal{F}$. If $\{x_t\}$ is a stochastic process with state space X, as usual, it will be called μ-measurable iff $x(\cdot, \cdot)$ is $(\mu \times P)$-measurable.

PROPOSITION 34. *Suppose that X is metric. Then if the process $\{x_t\}$ has a μ-measurable standard modification the function $\hat{x}_.$ is μ-measurable. Conversely, if $\hat{x}_.$ is μ-measurable, there is a naturally μ-measurable standard modification.*

This proposition is due to Yukiyosi Kawada [6] aside from the 'naturally.' It is easily deduced from Proposition 32 by exploiting the relations between Borel measurable functions and functions measurable with respect to the domain of a measure, or can be deduced directly, as Kawada did.

PROPOSITION 35. *If X is compact metric, the standard modification described in Proposition 34 can be made separable (or even right separable) in addition to the other stated properties.*

Proof. The proof of Proposition 33 together with the known relations between Borel and μ-measurability yields a sequence $\{v_{jj}\}$ such that $\lim_{j\to\infty} x[\phi_{v_{jj}}(t), \omega]$ exists for all (t, ω) and, if t is not in an exceptional Borel set B of μ-measure 0, the limit is $x(t, \omega)$ with probability 1. If t is not in B define $x_0(t, \omega)$ as the above limit. If t is in B any definition of $x_0(t, \omega)$ making $x_{0t} = x_t$ with probability 1 will yield an x_{0t} process which is a progressively μ-measurable standard modification of the x_t process. The standard separability argument yields a choice of x_{0t} making the process separable, or right separable if desired. (In the latter case ϕ_n must be chosen to make $\phi_n(t) \geqq t$.)

5. Examples. The following simple process may be used to furnish several examples alluded to in preceding sections. It will be described in an informal way using the terminology of [1], to which we refer for rigorous details.

There is a homogeneous Markov chain[6] $\{x_t, t \geqq 0\}$ on a probability space $(\Omega, \mathcal{F}, P)$ with three states $\{0, 1, 2\}$ having the following properties. The

[6] Also called "Markov chain with stationary transition probabilities."

mean sojourn time in each state is 1. The initial state is 0, upon exit from which there is a jump to state 1 or state 2 with probability 1/2 each. Upon exit from 1 there is a jump to 2, and vice versa. Each sample function takes the value 0 in a proper interval beginning at 0 and takes the values 1 and 2 (whichever comes first) thereafter in alternate intervals extending to infinity. In one version x_+ of the process every sample function is right continuous, in another version x_-, it is left continuous. Let α be the first entrance time into the state 1 and β that into the state 2. Both have the same density function, and $\alpha \wedge \beta = \gamma$ is the exit time from the state 0, with the density $e^{-t}\,dt$.

Let $\{\mathcal{F}_t, t \geqq 0\}$ be the natural family of $\{x_+(t), t \geqq 0\}$ or $\{x_-(t), t \geqq 0\}$; for each t let $\mathcal{F}_t^*$ be the smallest Borel field containing $\mathcal{F}_t$ and all sets of probability zero. Where no version of the above process is specified below, either x_+ or x_- will do.

Example 1. Relative to the family $\{\mathcal{F}_t, t \geqq 0\}$, α is optional but not strictly so. Since $P\{\alpha = t\} = 0$ for each t, α is strictly optional relative to $\{\mathcal{F}_t^*, t \geqq 0\}$.

Example 2. Relative to $\{\mathcal{F}_t^*\}$, we have $\mathcal{F}_{\alpha^*} = \mathcal{F}\{\alpha\}$, the Borel field generated by α alone. But

$$x(\alpha + 0) = \lim_{t \downarrow \alpha(\omega)} x(t, \omega) \in \mathcal{F}_{\alpha+} = \mathcal{F}_\alpha$$

and $P\{x(\alpha + 0) = 1\} = P\{x(\alpha + 0) = 2\} = 1/2$. Thus α is strictly optional but $\mathcal{F}_{\alpha^*} \neq \mathcal{F}_\alpha$.

Example 3. Relative to $\{\mathcal{F}_t\}$ or $\{\mathcal{F}_t^*\}$, we have

$$\mathcal{F}_{\gamma-} = \mathcal{F}_{(\alpha \wedge \beta)-} = \mathcal{F}\{\gamma\}$$

or the augmentation of $\mathcal{F}\{\gamma\}$ by all sets of probability zero. The set $\{\alpha = \beta\}$ is empty; the set $\{\alpha < \beta\}$ has probability 1/2 and is independent of the random variable γ. We have by (23),

$$\{\alpha < \beta\} \in \mathcal{F}_{\beta-}, \qquad \{\alpha < \beta\} = \Omega - \{\beta < \alpha\} \in \mathcal{F}_{\alpha-},$$

hence

$$\{\alpha < \beta\} \in \mathcal{F}_{\alpha-} \wedge \mathcal{F}_{\beta-}$$

but

$$\{\alpha < \beta\} \notin \mathcal{F}_{(\alpha \wedge \beta)-}.$$

Example 4. For the x_- version, $x_{\alpha \wedge t} \equiv 0$ for every $t \geqq 0$. Hence

$$\alpha \notin \mathcal{F}\{x_{\alpha \wedge t}, t \geqq 0\}.$$

Example 5. Let $\Delta \notin \mathcal{F}_{\gamma+0}$, e.g., $\Delta = \{\alpha > 1\}$; and define

$$\tilde{x}(t,\omega) = \begin{cases} x(t,\omega), & \text{if } t \neq \gamma(\omega); \\ x(\gamma(\omega)-0,\omega), & \text{if } t = \gamma(\omega), \omega \in \Delta; \\ x(\gamma(\omega)+0,\omega), & \text{if } t = \gamma(\omega), \omega \in \Omega - \Delta. \end{cases}$$

Then $\{\tilde{x}_t\}$ is a (separable) standard modification of $\{x_t\}$ and so has the same augmented natural family $\{\mathcal{F}_t^*\}$ as $\{x_t\}$. We have $\tilde{x}_\gamma \notin \mathcal{F}_{\gamma+0}^*$ since

$$\{\tilde{x}_\gamma = 0\} = \Delta \notin \mathcal{F}_{\gamma+0}^*.$$

Example 6. For the x_+ process, the set $\{\alpha < \beta\}$ belongs to $\mathcal{F}_{\alpha*}{}^0$ but not to $\mathcal{F}_{\alpha-}{}^0$. The process $\{x_t^*\}$ is not separable, but an obvious discrete parameter analogue serves the same purpose and eliminates the question of separability.

Example 7. For the minimal chain studied in [1; § II.19] and [2], we have, if τ_n is the n-th jump,

$$\bigvee_{n=1}^{\infty} \tau_n = \lim \tau_n = \tau$$

where τ is the "first infinity." The random variable $x(\tau+0)$ does not belong to $\bigvee_{n=1}^{\infty} \mathcal{F}_{\tau_n+}$ but belongs to $\mathcal{F}_{\tau+}$. See Theorem 4.4 of [2].

Acknowledgment. A number of results similar to those in this paper were obtained independently by P. A. Meyer and will appear in his forthcoming book.

STANFORD UNIVERSITY
AND
UNIVERSITY OF ILLINOIS.

REFERENCES.

[1] K. L. Chung, *Markov chains with stationary transition probabilities*, Springer, 1960.

[2] ———, "On the boundary theory for Markov chains," *Acta Mathematica*, vol. 110 (1963), pp. 19-77.

[3] J. L. Doob, *Stochastic processes*, Wiley, 1953.

[4] E. B. Dynkin, *Foundations of the theory of Markov processes* (in Russian), Moscow, 1963.

[5] P. R. Halmos, *Measure theory*, Van Nostrand, 1950.

[6] Y. Kawada, "On measurable stochastic process," *Proceedings of the Physics and Mathematics Society of Japan*, vol. 23 (1941), pp. 511-527.

Acta Math. **115** (1966) 111–163.
With kind permission of Springer Science and Business Media.

ON THE BOUNDARY THEORY FOR MARKOV CHAINS. II

BY

KAI LAI CHUNG

Stanford University, Stanford, California, U.S.A.(1)

§ 11. Introduction

This paper may be regarded as a new and fairly self-contained one attached to §§ 2–4 of [I].(2) These sections are entitled "Terminology and notation", "The boundary" and "Fundamental theorems" respectively. The rest of [I] is either contained in a more general treatment (§ 5, § 9 and parts of § 6), or may be set aside as special cases under additional hypotheses (parts of § 6, § 7 and § 8). In particular the whole idea of "dual boundary" is dispensed with here, though this is not to say it should be abandoned forever. References to [I] beyond § 4 will be pinpointed.

In sum, the case of a finite number of passable atomic boundary points (briefly: "exits") will be settled here. Namely: all homogeneous Markov chains satisfying Assumptions A and B′ [I; p. 25 and p. 50] will be completely analyzed, with regard to the stochastic behavior of the sample functions as well as the analytical structure of transition probabilities, in fact both at the same time. To be exact, it will also be assumed that:

ASSUMPTION C_1. *All Φ-recurrent states are merged into one absorbing state.*

ASSUMPTION D. *All exits are distinguishable.*

It is important to note the difference between C_1 above and the erstwhile Assumption C [I; p. 47] which would require the absence of any $\prod$-recurrent state and is a serious restriction. On the contrary, conditions C_1 and D may be justly regarded as unessential for the boundary theory; see respectively the discussion at the end of § 15 here and on p. 38 of [I].

A culminating result of the theory has been that of "complete construction", origi-

(1) This research is supported in part by the Office of Scientific Research of the United States Air Force.

(2) References in roman capitals are listed at the end of the paper; references [1] to [14] are to be found at the end of [I].

nated with Feller [9], developed through Neveu [11],(1) and established first by David Williams [VIII] under substantially the same conditions as Theorem 16.1 here, in a somewhat different form and with a totally different, purely analytic method. This result falls into two parts. The first, to be called "decomposition" here, is an analysis of the basic transition matrix (or its Laplace transform) by decomposing it into several components: the exits z, the entrances η, and a connecting matrix M (see § 16 for these symbols). The second, to be called "construction", consists in showing that any such transition matrix can be so put together from similar components chosen rather arbitrarily but subject to certain analytical conditions. Now it should be apparent that the decomposition is in general not unique without further conditions on the choices (such as choosing both the exits and entrances to be "extreme bases" as in Feller's case), and without uniqueness the two parts of the theorem are not really in direct correspondence. Thus, if a process is constructed and then decomposed, the original components used in the construction are not necessarily thereby retrieved. To put it in another way, in an arbitrary construction the various components need not have the meanings attached to those in a meaningful decomposition, although the corresponding (and cognate) parts look quite like each other formally. To see that this question is not an academic one, consider the following problem: from a given process, to construct a new one by stopping it at certain specified exits (see § 18 for a precise formulation and solution). Obviously, this problem cannot be solved by another construction using only the "through" exists, because the corresponding entrances can no longer be chosen arbitrarily. Rather, one must begin with a correct decomposition of the original process, and then shut off the properly identified entrances.

Such a decomposition will be called "canonical" and it will be derived by the most natural probabilistic considerations.(2) As a matter of fact, the canonical form conceals a more fundamental resolution into elements which are simpler to define and easier to use. These are the probabilities ϱ^a and F^{ab} introduced and studied in § 14. Each ϱ^a is then linked to an entrance law η^a through a measure E^a (Theorem 14.4), the meaning of which is given in § 17. The canonical decomposition itself, in these two stages, is given in Theorem 15.2.

For a fuller understanding of the stochastic as well as analytic structure of the process, however, we must consider a third problem, that of "identification", to be taken up in

---

(1) It should be pointed out that Neveu's results do not seem to include Feller's since Theorem 4.2.1 of [11] requires, besides "absolute dominance", also e.g. that the cone of entrance laws *relative to* Π (*rather than* Φ in our notation) be of finite dimension. This is a quite different type of assumption from those made by all the other authors.

(2) Observe, *inter alia*, that in the form given here the substochastic case becomes an easy extension of the stochastic one, and that each entrance law is generated by an entrance sequence (excessive measure relative to Φ).

§§ 17–18. Here the major results are given in Theorems 17.1 and 17.3. In contrast to some prior developments, these no longer appear to be "intuitively" obvious", and yet they depend crucially on the set-theoretic properties of "boundary times" given in § 12. One is convinced by the amount of detection needed to identify such simple quantities as ϱ^a and F^{ab} that herein lies indeed the strength of the probabilistic versus the analytico-algebraic method.

In § 18 algebraic transformations between different decompositions are established and an example is given to clarify the question of construction discussed above. In § 19 some consequences of the main results are specified ending with a full description of the sample functions of the process in terms of all the quantities introduced in this paper.

§ 12. Classification of boundary atoms

In this section properties of individual passable atomic boundary points, to be called "boundary atoms" for brevity's sake from now on, will be discussed. No hypothesis beyond Assumption A (p. 25 of [I]) and the existence of such an atom is needed. These atoms will be denoted simply by $a, b, \ldots$ instead of $\infty^a, \infty^b, \ldots$, and the set by $\mathbf{A}$. Correspondingly we shall write $S^a(\omega)$ for $S_{\infty^a}(\omega)$ (p. 38 of [I]) and "$x(t)=a$" for "$t\in S^a(\omega)$" or "x reaches the boundary atom a at time t". Furthermore we shall define $\mathbf{P}^a\{\ldots\}$ by the condition that for every $t_\nu\in\mathbf{T}, j_\nu\in\mathbf{I}, 1\leqslant\nu\leqslant n$, we have

$$\mathbf{P}^a\{x(t_\nu)=j_\nu,\ 1\leqslant\nu\leqslant n\}=\mathbf{P}\{x^a(t_\nu)=j_\nu\,|\,\Delta^a\},$$

where the right-hand side is defined in the last paragraph on p. 34 of [I]. This uniquely defines a probability measure on the Borel field $\mathfrak{F}^0$ generated by the Markov chain $\{x_t, t\in\mathbf{T}\}$. It is "the conditional probability when the process starts at a"; note the analogy with the usual $\mathbf{P}_i\{\ldots\}$ for $i\in\mathbf{I}$ (p. 24 of [I]). Similarly for conditional expectation $\mathbf{E}^a\{\ldots\}$.

"For a.e. ω" will mean for every ω except a set N in $\mathfrak{F}^0$ such that $\mathbf{P}_i(N)=0$ for each i in $\mathbf{I}$. It will then follow that we have also $\mathbf{P}^a(N)=0$ for each a in $\mathbf{A}$.

Let us define, for $s\geqslant 0$:

$$\alpha_s(\omega)=\inf\{t:t>s:x(t)\in\mathbf{A}\}, \tag{12.1}$$

where the inf is taken to be $+\infty$ if the set is empty; and

$$\alpha(\omega)=\alpha_0(\omega).$$

Thus α_s is the "first time after s when the boundary is reached"; it is an optional random variable; and α is the τ on p. 25 of [I], the letter τ (with subscripts) being reserved for a general time random variable in this paper. Next, we define

$$\begin{aligned}\alpha^a(\omega) &= \inf\{t: t>0,\, x(t)=a\};\\ K_i^a(t) &= \mathbf{P}_i\{\alpha^a \leqslant t\};\\ K^{ab}(t) &= \mathbf{P}^a\{\alpha^b \leqslant t\}.\end{aligned} \tag{12.2}$$

Thus K^b and K^{ab} are the first entrance time distributions into b, starting at i and a respectively. Extending a familiar notation in [1], we write $i \rightsquigarrow a$ iff $K_i^a(\infty)>0$; $a \rightsquigarrow b$ iff $K^{ab}(\infty)>0$;(1) and $a \rightsquigarrow i$ iff $i \in \mathbf{I}^a$, i.e., iff $\xi_i^a(t) \not\equiv 0$ (see pp. 34–5 of [I] for notation). According to § 10 of [I], $\xi_i^a(\cdot)$ is either identically zero or never zero; similar properties hold for K_i^a and K^{ab} but the full strength of this result will not be needed. It is now possible to define the relation $\leftrightsquigarrow$, called "communication", between any two elements of $\mathbf{I} \cup \mathbf{A}$ in the obvious way and deduce the usual classification (see [1; § II.10]). We shall give only the few propositions that will be needed later.

Definition 12.1. The boundary atom a is called *recurrent* if

$$\mathbf{P}^a\{S^a(\omega) \text{ is an unbounded set }\}=1;$$

otherwise it is called *nonrecurrent.*

THEOREM 12.1. *If a is recurrent, then for every $\delta>0$:*

$$\mathbf{P}^a\{S^a(\omega)\cap(\delta,\infty)\neq\emptyset\}=1.$$

Conversely if there exists a $\delta>0$ for which (12.2) *is true, then a is recurrent.*

Proof. Clearly Definition 12.1 implies

$$\mathbf{P}^a\{\forall\delta>0: S^a(\omega)\cap(\delta,\infty)\neq\emptyset\}=1$$

which implies the first assertion. To prove the converse, let $\tau_0(\omega)\equiv 0$ and for $n\geqslant 0$ define

$$\tau_{n+1}(\omega)=\inf\{t: t>\tau_n(\omega)+\delta,\, x(t)=a\}.$$

By the strong Markov property [1; Theorem II.9.3] the fields $\mathfrak{F}_{\tau_n+\delta}$ and $\mathfrak{F}'_{\tau_n+\delta}$ are independent conditioned on $x(\tau_n+\delta)$ which is in $\mathbf{I}$ with probability one. Also τ_{n+1} is measurable $\mathfrak{F}'_{\tau_n+\delta}$ and is the first entrance time into a in the post-$(\tau_n+\delta)$ process. Hence we have by the Strong Markov property:(2)

(1) Note that this is not the same definition as on p. 43 of [I].

(2) This will be used so often that we cannot mention it every time, but it must be remembered that we are invoking here the form for boundary entrances as given in [I; Theorem 4.4], rather than the usual form as given in (1; Theorem II.9.3). We will indicate this by using the capital S for the former and the small s for the latter.

$$\mathbf{P}\{\tau_{n+1} < \infty \,|\, \tau_1, \ldots, \tau_n\} = \mathbf{P}\{\tau_{n+1} < \infty \,|\, x(\tau_n)\} = \mathbf{P}^a\{S^a(\omega) \cap (\delta, \infty) \neq \emptyset\} = 1.$$

Consequently all τ_n are finite, $\tau_n \uparrow \infty$ and S^a is unbounded with probability one.

THEOREM 12.2. *If a is recurrent, then $\mathbf{I}^a$ is a $\prod$-recurrent class and for every $i \in \mathbf{I}^a$ we have $K_i^a(\infty) = 1$. Conversely, if there exists a $\prod$-recurrent state i in $\mathbf{I}^a$ such that $i \rightsquigarrow a$ then a is recurrent.*

Proof. Let $i \in \mathbf{I}^a$, then $\xi_i^a(t)$ is positive from a certain t on (in fact for all $t > 0$ by § 10 of [I]). Hence there exists $\delta > 0$ such that

$$\int_0^\delta \xi_i^a(t)\, dt > 0. \tag{12.3}$$

Using the sequence $\{\tau_n\}$ defined above, we have

$$\int_0^\infty \xi_i^a(t)\, dt \geqslant \sum_{n=1}^\infty \mathbf{E}\{\mu[S_i \cap (\tau_n, \tau_{n+1})]\} \geqslant \sum_{n=1}^\infty \int_0^\delta \xi_i^a(t)\, dt = +\infty.$$

Furthermore by Theorem 12.1, we have

$$0 = \mathbf{P}^a\{S^a \cap (\delta, \infty) = \emptyset\} \geqslant \xi_i^a(\delta)\,[1 - K_i^a(\infty)].$$

But (12.3) implies that $\xi_i^a(\delta) > 0$ by the first sentence of the proof, hence $K_i^a(\infty) = 1$. Since $p_{ii}(t) \geqslant \int_0^t \xi_i^a(t-s)\, dK_i^a(s)$ it follows that

$$\int_0^\infty p_{ii}(t)\, dt \geqslant K_i^a(\infty) \int_0^\infty \xi_i^a(t)\, dt = +\infty;$$

hence i is recurrent [1; Theorem II.10.4]. For each j in $\mathbf{I}^a$, we have $i \rightsquigarrow a \rightsquigarrow j$; hence $\mathbf{I}^a$ is one $\prod$-recurrent class.

Conversely, let $i \in \mathbf{I}^a$ and i be $\prod$-recurrent. Then

$$\mathbf{P}^a\{S_i \text{ is an unbounded set}\} = 1. \tag{12.4}$$

If $K_i^a(\infty) > 0$, there exist $\delta > 0$ such that $K_i^a(\delta) > 0$. Define a new sequence $\{\tau_n'\}$ as before but with "a" replaced by "i". It follows from (12.4) that all τ_n' are finite and $\tau_n' \uparrow \infty$ with probability one. We have

$$\mathbf{P}^a\{S^a \cap (\tau_n', \tau_{n+1}') \neq \emptyset\} \geqslant K_i^a(\delta)$$

and the events Λ_n in the $\{\cdots\}$ above are independent by the strongest Markov property [1; Theorem II.9.5] applied to the τ's. Hence by the Borel–Cantelli lemma, infinitely many of the Λ's occur and so S^a is unbounded with probability one.

COROLLARY. *If a is nonrecurrent, and $i \in \mathbf{I}^a$, then either i is $\prod$-nonrecurrent or $i \not\rightsquigarrow a$ (i.e., the negation of $i \rightsquigarrow a$). Conversely, if there exists an i in $\mathbf{I}^a$ such that either $i \not\rightsquigarrow a$ or i is $\prod$-nonrecurrent, then a is nonrecurrent.*

However, $\mathbf{I}^a$ may contain more than one distinct class. Let us also observe that if i is $\prod$-recurrent and $i \not\rightsquigarrow \mathbf{A}$ (i.e., $\forall a \in A: i \not\rightsquigarrow a$), then i must be Φ-recurrent. For if $i \not\rightsquigarrow A$, then $\forall j \in \mathbf{I}$ we have $p_{ij}(\cdot) \equiv f_{ij}(\cdot)$ and consequently $\prod$-recurrence of i implies its Φ-recurrence. Conversely, if i is Φ-recurrent then it is certainly $\prod$-recurrent and $i \not\rightsquigarrow \mathbf{A}$ by Theorem 3.2 of [I].

Definition 12.2. The boundary atom a is called *sticky* iff

$$\mathbf{P}^a\{\forall \delta > 0: S^a \cap (0, \delta) \neq \emptyset\} = 1; \tag{12.5}$$

otherwise it is called *nonsticky* (it will follow after Theorem 12.5 that the probability above is then equal to 0).

We begin with a simple observation valid for every a.

THEOREM 12.3. *For each a and a.e. ω, $S^a(\omega)$ is a countable set.*

Proof. The definition of "reaching the boundary" (pp. 28–29 of [I]) entails that if $t \in S^a(\omega)$, then there exists $\delta > 0$ such that $(t-\delta, t) \notin S^a(\omega)$. Hence every point in $S^a(\omega)$ is isolated on the left, and the theorem follows from a well-known property of the real line.

THEOREM 12.4. *If a is sticky, then for a.e. ω, $S^a(\omega)$ is dense in itself.*

Remark. In view of the preceding proof this means: for each t in $S^a(\omega)$ and $\delta > 0$, we have $(t, t+\delta) \cap S^a(\omega) \neq \emptyset$. For $t = 0$ this reduces to (12.5).

Proof. For each real r and for a.e. ω for which $\alpha_r(\omega) < \infty$, we have by (12.5) and the Strong Markov property:

$$\forall \delta > 0: \quad S^a(\omega) \cap (\alpha_r(\omega), \alpha_r(\omega) + \delta) \neq \emptyset.$$

Hence for a.e. ω this is even true for all $\alpha_r(\omega)$ with rational values of r, simultaneously. Since every point in $S^a(\omega)$ is such an $\alpha_r(\omega)$ by definition, Theorem 12.4 is proved.

COROLLARY. *If a is sticky and*

$$\gamma(t, \omega) = \sup [S^a(\omega) \cap (0, t)],$$

then $\gamma(t, \omega) \in \overline{S^a(\omega)} - S^a(\omega)$ where $\overline{S^a}$ is the Euclidean closure of S^a.

Proof. By definition, $\gamma(t, \omega) \in \overline{S^a(\omega)}$; by the Remark above $\gamma(t, \omega) \notin S^a(\omega)$.

The next result is a zero-or-one law for the notion of stickiness. It may be remarked that the so-called zero-or-one law in the theory of Hunt processes is trivially false for Markov chains in general.

THEOREM 12.5. *If a is nonsticky, then for a.e. ω, $S^a(\omega)$ does not have a finite point of accumulation. In particular, the probability in* (12.5) *is equal to zero.*

Proof. Since (12.5) does not hold, there exist $\delta > 0$ and $\varepsilon > 0$ such that

$$\mathbf{P}^a\{S^a \cap (0, \delta) \neq \emptyset\} = 1 - \varepsilon. \tag{12.6}$$

Suppose the theorem false and let Λ_0 be a set of positive probability such that if $\omega \in \Lambda_0$ then $S^a(\omega)$ has a finite point of accumulation. Then there exist $t \geq 0$, and a subset Λ_1 of Λ_0 with positive probability such that if $\omega \in \Lambda_1$ then $S^a(\omega)$ has a point of accumulation in $(t, t+\delta)$. This implies that $S^a(\omega) \cap (t, t+\delta)$ is an infinite set. Now for each $m \geq 1$, let

$$t < \tau_{m1}(\omega) < \tau_{m2}(\omega) < \ldots < \tau_{mN}(\omega) < t + \delta$$

be all the successive distinct members of the set $\{\alpha_{n/m}(\omega),\ n \geq 0\}$ defined in (12.1), where $N = N(m, \omega)$ is a nonnegative integer. For each ω in Λ_1, we have $\lim_{m\to\infty} N(m, \omega) = +\infty$. Hence given N_0, there exist m_0 and $\Lambda_2 \subset \Lambda_1$, with $2P(\Lambda_2) \geq P(\Lambda_1)$, such that

$$\forall \omega \in \Lambda_2: \quad N(m_0, \omega) > N_0.$$

Applying the Strong Markov property to $\tau_{m_0 1}, \ldots, \tau_{m_0 N_0}$, we obtain by (12.6):

$$\mathbf{P}\{\tau_{m_0, n+1} - \tau_{m_0 n} < \delta \,|\, \tau_{m_0 1} < \ldots < \tau_{m_0 n} < \infty\} \leq P^a\{S^a \cap (0, \delta) \neq \emptyset\} = 1 - \varepsilon.$$

It follows that

$$\mathbf{P}(\Lambda_1) \leq 2\mathbf{P}(\Lambda_2) \leq 2(1-\varepsilon)^{N_0}.$$

Since N_0 is arbitrary, $\mathbf{P}(\Lambda_1) = 0$. This is a contradiction that proves the theorem.

§ 13. Exit and cntrance sequcnces and laws

This short section contains several more-or-less known propositions in the forms to be needed later.

Given the countable index set $\mathbf{I}$, let $\mathfrak{M}(\mathbf{I})$ be the space of measures on $\mathbf{I}$, namely al sequences of nonnegative finite real numbers index by $\mathbf{I}$.

Definition 13.1. Given a standard substochastic transition matrix function $\Psi(\cdot)$ on $\mathbf{I} \times \mathbf{I}$, an entrance (exit) sequence e relative to Ψ is an element of $\mathfrak{M}(\mathbf{I})$ satisfying

$$\forall t \geq 0: \quad e \geq e\Psi(t) \quad [e \geq \Psi(t)e].$$

Here if $e=\{e_i\}$ and $(\Psi(t))=(\psi_{ij}(t))$, $e\Psi(t)$ is the element of $\mathcal{M}(\mathrm{I})$ whose j-component $[e\Psi(t)]_j$ is $\sum_i e_i\psi_{ij}(t)$ [$\Psi(t)e$ is the element whose i-component is $\sum_j \psi_{ij}(t)e_j$]; and the inequality is taken component-wise. We shall use this type of "vector notation" when confusion is unlikely; but since both subscripts and superscripts will appear as possible components of vectors we shall revert to an explicit notation whenever in doubt.

Since $\Psi(\cdot)$ is standard, we have $\lim_{t\downarrow 0}\psi_{ii}(t)=1$ for every i, from which it follows easily that

$$e=\lim_{t\downarrow 0} e\Psi(t) \quad [e=\lim_{t\downarrow 0}\Psi(t)e],$$

so that e is "excessive" in Hunt's usage. It is easy to prove that relative to the minimal solution $\Phi(\cdot)$ (see p. 23 of [I]), e is an entrance (exit) sequence if and only if

$$eQ\leqslant 0 \quad [Qe\leqslant 0],$$

where Q is the initial derivative matrix.

We shall state the next two theorems for the entrance case only since the exit case is entirely similar.

Definition 13.2. An *entrance law* relative to Ψ is a one-parameter family $\eta(\cdot)=\{\eta(s),\ s\geqslant 0\}$ of elements of $\mathcal{M}(\mathrm{I})$ satisfying the functional equation:

$$\forall s>0,\ t\geqslant 0:\quad \eta(s)\Psi(t)=\eta(s+t). \tag{13.1}$$

By [3; Lemma 1], for each j in I, $\eta_j(\cdot)$ is continuous in $[0,\infty)$.

THEOREM 13.1. *If* e *and* Ψ *are as in Definition* 13.1, *there exists an entrance law* $\eta(\cdot)$ *relative to* Ψ *such that*

$$e-e\Psi(t)=\int_0^t \eta(s)\,ds. \tag{13.2}$$

Remark. We shall say that the entrance sequence e generates the entrance law $\eta(\cdot)$.

Proof. This is nothing but a general form of Theorem 6.2 of [I] proved in the same way, but a sketch will be given. Let

$$H(t)\overset{\mathrm{def}}{=}e-e\Psi(t). \tag{13.3}$$

It is clear that $H(\cdot)\nearrow$ and the semigroup property $\Psi(s)\Psi(t)=\Psi(s+t)$ implies

$$H(s+t)-H(t)=H(s)\Psi(t). \tag{13.4}$$

By a basic lemma [3; Lemma 2], H has a continuous derivative η so that

$$H(t)=\int_0^t \eta(s)\,ds \tag{13.5}$$

which is (13.2); moreover the equation (13.4) may be differentiated with respect to s, yielding (13.1).

Theorem 13.2. *Let $\Psi(\cdot)$ be as before and let $\{e(s),\ s>0\}$ be a one-parameter family of elements of $\mathfrak{M}(\mathrm{I})$ satisfying*

$$e(s)\Psi(t)\leqslant e(s+t), \tag{13.6}$$

$$e\overset{\text{def}}{=}\int_0^\infty e(s)\,ds<\infty. \tag{13.7}$$

Then e is an entrance sequence relative to Ψ, and

$$\lim_{t\uparrow\infty} e\Psi(t)=0, \tag{13.8}$$

or equivalently

$$e=\int_0^\infty \eta(s)\,ds, \tag{13.9}$$

where η is the entrance law generated by e according to Theorem 13.1.

Proof. Integrating (13.6) over s in $(0,\infty)$ we obtain

$$e\Psi(t)\leqslant\int_0^\infty e(s+t)\,ds=e-\int_0^t e(s)\,ds\leqslant e.$$

Hence e is an entrance sequence as asserted, and also (13.8) is true by (13.7). The equivalence of (13.8) and (13.9) is obvious from (13.2).

It is instructive to compare the results above with the standard potential theory argument, according to which we should write

$$\int_0^\infty [e-e\Psi(t)]\,\Psi(u)\,du=\int_0^\infty H(t)\,\Psi(u)\,du=\int_0^\infty [H(t+u)-H(u)]\,du=\int_0^t [H(\infty)-H(u)]\,du.$$

Therefore in particular

$$\lim_{t\downarrow 0}\int_0^\infty \frac{1}{t}\,[H(t+u)-H(u)]\,du=H(\infty)-H(0)=e.$$

Under our conditions it is permitted to interchange the limit and integration above which then becomes (13.9). However, if the same interchange is made a little earlier in

$$\lim_{t\downarrow 0}\frac{1}{t}\int_0^\infty H(t)\,\Psi(u)\,du,$$

the result is
$$\int_0^\infty \eta(0)\,\Psi(u)\,du$$

which is in general strictly less than e.

THEOREM 13.3. *Let*
$$z \overset{\text{def}}{=} \lim_{t\uparrow\infty} [I-\Psi(t)]\,1; \tag{13.10}$$

then z is an exit sequence relative to Ψ. If e is any exit sequence, then
$$e \leqslant z \tag{13.11}$$

if and only if
$$e \leqslant 1 \quad \textit{and} \quad \lim_{t\uparrow\infty} \Psi(t)\,e = 0. \tag{13.12}$$

Remark. By the analogue of Theorem 13.2, the second condition in (13.12) is equivalent to that $e=\int_0^\infty e(s)\,ds$ where $e(\cdot)$ is the exit law generated by e.

Proof. It is clear that $z\leqslant 1$ and $\lim_{t\uparrow\infty}\Psi(t)z=0$, hence (13.11) implies (13.12). Conversely if (13.12) holds, then
$$[I-\Psi(t)]\,e \leqslant [I-\Psi(t)]\,1$$

and letting $t\uparrow\infty$ we obtain (3.11). Q.e.d.

When Ψ is the minimal solution Φ, then z is $L(\infty)$. From the probabilistic point of view there is an obvious choice of exit sequences relative to Φ. They are the $L^a(\infty)$, $a\in \mathbf{A}$ studied in § 4 of [I]. It is the entrance sequences that have to be discovered and this will be done in Theorem 14.3 below.

§ 14. The basic quantities

From here on Assumptions A and B′ will be in force throughout the paper. In this section we introduce the new notions which make the present approach possible. The underlying idea is simple enough: to study the succession of boundary atoms in a sample function, viewing these as "banners" (superscripts from $\mathbf{A}$) under which the ordinary states (subscripts from $\mathbf{I}$) line up. If all boundary atoms are nonsticky, this is easily carried out, has essentially been done in § 5 of [I] and will be reviewed in § 19. For a sticky atom the important thing is to concentrate on the *change* of banners so that beginning at one of them the sample function is followed through until a new one appears, if ever. Now it turns out that each portion between change of banners "possesses finite potentials" (Theorem 14.2) so that it can be sufficiently well isolated and analyzed before the portions are pieced together. Special attention must be paid to the case when the banner does not change and

when it changes suddenly. The first can be treated separately as "traps"; the second produces delicate effects which will be stressed at the appropriate places.

Definition 14.1. For each a in $\mathbf{A}$ let us define a new optional random variable as follows:

$$\beta^a(\omega)=\inf_{b\neq a}\alpha^b(\omega)=\inf\{t:t>0,\ x(t)\in\mathbf{A}-\{a\}\},$$

where α^b is defined in (12.2). Thus starting at a, β^a is the "first time for change (of banners)".

Definition 14.2. For each $a\in\mathbf{A}$, $b\in\mathbf{A}$, $a\neq b$ and $t\geqslant 0$:

$$\varrho_j^a(t)\overset{\text{def}}{=}\mathbf{P}^a\{\beta^a>t;\ x(t)=j\},$$

$$F^{ab}(t)\overset{\text{def}}{=}\mathbf{P}^a\{\alpha^b=\beta^a\leqslant t\}.$$

Thus $\varrho_j^a(t)$ is the probability, starting at a, that no change of banner has occurred up to time t and that at this time state j appears under the initial banner a; $F^{ab}(t)$ is the probability that a change of banner has occurred before or on time t and that the change is to b (regardless what banner is flying at t).

We have the obvious relations, if $t>0$:

$$\varrho_*^a(t)\overset{\text{def}}{=}\sum_i\varrho_i^a(t)=\mathbf{P}^a\{\beta^a>t\},\quad \sum_{b\neq a}F^{ab}(t)=\mathbf{P}^a\{\beta^a\leqslant t\};\tag{14.1}$$

$$\varrho_*^a(t)+\sum_{b\neq a}F^{ab}(t)=1.\tag{14.2}$$

Note that $\varrho_j^a(0)=\varrho_j^a(0+)$ but $\varrho_*^a(0)\leqslant\varrho_*^a(0+)$ in general since $\sum_j\mathbf{P}^a\{x(0)=j\}\leqslant 1$. The limits of (14.2) at either end of $(0,\infty)$ are important:

$$\sum_{b\neq a}F^{ab}(0)=1-\lim_{t\downarrow 0}\uparrow\varrho_*^a(t)=1-\varrho_*^a(0+);\tag{14.3}$$

$$\sum_{b\neq a}F^{ab}(\infty)=1-\lim_{t\uparrow\infty}\downarrow\varrho_*^a(t)=1-\varrho_*^a(\infty).\tag{14.4}$$

Definition 14.3. The boundary atom is called *ephemeral* iff $\varrho_*^a(0+)=0$; it is called a *trap* iff $\varrho_*^a(\infty)=1$.

It is clear how we can eliminate each ephemeral boundary atom by splitting it into others, but this will not be necessary.

THEOREM 14.1. *If a is sticky and distinguishable from any other boundary atom, then*

$$\sum_{b\neq a}F^{ab}(0)=0.\tag{14.5}$$

If b is nonsticky, then

$$\sum_{a \neq b} F^{ab}(0) = 0. \tag{14.6}$$

Proof. It follows from Definition 14.2 that

$$F^{ab}(0) = \mathbf{P}\{\forall \delta > 0 : S^b \cap (0; \delta) \neq \emptyset\}. \tag{14.7}$$

If a is sticky then by definition, 0 is an accumulation point of $S^a(\omega)$ for almost every ω. Hence by Theorem 4.6 of [I], for almost no ω can 0 be an accumulation point of any $S^b(\omega)$ where b is distinguishable from a. This means $F^{ab}(0) = 0$ by (14.7), and the first assertion of the theorem follows. Next suppose $F^{ab}(0) > 0$; choose i with $L_i^a(\infty) > 0$, then by the Strong Markov property,

$$\mathbf{P}_i\{\forall \delta > 0 : S^b \cap (\alpha^a, \alpha^a + \delta) \neq \emptyset\} \geqslant L_i^a(\infty) F^{ab}(0) > 0.$$

Thus S^b has a finite accumulation point α^a with positive probability and so b must be sticky by Theorem 12.5. The second assertion of the theorem follows. Q.e.d.

It is clear from the meaning of ϱ^a that (see p. 40 of [I] for ζ^a):

$$\zeta^a(t) \leqslant \varrho^a(t) \leqslant \xi^a(t).$$

Now let us put

$$p_{ij}^a(t) \stackrel{\text{def}}{=} f_{ij}(t) + \int_0^t l_i^a(u)\, \varrho_i^a(t-u)\, du. \tag{14.8}$$

By the Strong Markov property, this is the probability of transition when the boundary set $\mathbf{A} - \{a\}$ is taboo; namely when the process is stopped at all boundary atoms except a. Let us call this stopped process, completed as usual by a new absorbing state θ, the a-process (see [1; p. 244 ff.]). It is clear that

$$\prod{}^a \stackrel{\text{def}}{=} (p_{ij}^a(\cdot)), \quad (i, j) \in \mathbf{I}^a \times \mathbf{I}^a,$$

is a substochastic transition matrix (function) whose stochastic completion is the transition matrix of the a-process and that

$$\Phi \leqslant \prod{}^a \leqslant \prod. \tag{14.9}$$

Moreover, if the process is initially at a it is an open Markov chain whose absolute distribution is $\{\varrho^a(t), t > 0\}$. This fact is expressed by the following functional equation

$$\varrho^a(s) \prod{}^a(t) = \varrho^a(s+t); \tag{14.10}$$

which is to be compared with

$$\zeta^a(s)\Phi(t) = \zeta^a(s+t), \qquad \xi^a(s)\prod(t) = \xi^a(s+t).$$

Thus we have interposed, for each a, a new process between the minimal Φ and the maximal $\prod$. Successive interposition will lead from Φ to $\prod$, but that will not be necessary.

The continuity of each $\varrho_j^a(\cdot)$ follows from (14.10) by [3; Lemma 1]. Furthermore each $\varrho_j^a(\cdot)$ has the property of being either identically zero or never zero in $(0,\infty)$, by [I; § 10].

If a is a trap, then clearly $\varrho^a(\cdot)\equiv\xi^a(\cdot)$. If a is not a trap, the following result is essential (see p. 29 of [I] for Z).

Theorem 14.2. *If a is not a trap, and $i\in \mathbf{I}-Z$, then*

$$r_i^a \stackrel{\text{def}}{=} \int_0^\infty \varrho_i^a(t)\,dt < \infty. \tag{14.11}$$

Proof. If $i\in\mathbf{I}-Z$, then $i\rightsquigarrow \mathbf{A}$. Since a is not a trap, $a\rightsquigarrow \mathbf{A}-\{a\}$ so that in any case $i\rightsquigarrow\mathbf{A}-\{a\}$. Hence there exists $b\in\mathbf{A}-\{a\}$, and $h>0$ with $L_i^b(h)>0$. Now

$$\varrho_i^a(nh)\,L_i^b(h)\leqslant \mathbf{P}^a\{nh<\beta^a\leqslant(n+1)\,h\}$$

and so
$$\sum_{n=1}^\infty \varrho_i^a(nh)\leqslant [L_i^b(h)]^{-1}<\infty.$$

Since
$$\varrho_i^a(t)\,f_{ii}(h)\leqslant\varrho_i^a(t+h)$$

and $\varrho_i^a(\cdot)$ is continuous, we have

$$\max_{nh\leqslant t\leqslant(n+1)h}\varrho_i^a(t)\leqslant[\min_{0\leqslant t\leqslant h}f_{ii}(t)]^{-1}\,\varrho_i^a(nh+h).$$

Consequently
$$r_i^a=\int_0^\infty\varrho_i^a(t)\,dt\leqslant[\min_{0\leqslant t\leqslant h}f_{ii}(t)]^{-1}\sum_{n=0}^\infty\varrho_i^a(nh+h)<\infty$$

and Theorem 14.2 is proved.

The next step is quite similar to the handling of ξ in the one-exit nonrecurrent case given in § 9 of [I]. The present extension to the general situation is made possible by the interposition of ϱ which behaves as "nonrecurrent" in the sense of the preceding theorem. Instead of Assumption C of [I; p. 47] the following assumption will be made from now on.

Assumption C_0. *There are no Φ-recurrent states.*

This will be slightly liberalized towards the end of § 15 to include the case of substochastic $\prod$, but it is not entirely dispensable without complicating the later results.

Theorem 14.3. *For each boundary atom a, there exists an entrance sequence e^a relative to Φ having the property that*

$$\lim_{t\uparrow\infty} e^a\Phi(t)=0.$$

Equivalently, if $\eta^a(\cdot)$ denotes the entrance law generated by e^a (see Theorem 13.1), we have

$$e^a = \int_0^\infty \eta^a(s)\, ds. \tag{14.12}$$

Proof. Three cases will be considered though the first two may be combined.

Case 1. a is not a trap. Then by (14.9) and (14.10),

$$\varrho^a(s)\Phi(t) \leqslant \varrho^a(s)\textstyle\prod^a(t) = \varrho^a(s+t)$$

and (14.11) is true. Hence we may take Ψ to be Φ, $e(s)$ to be $\varrho^a(s)$, e to be r^a and $\eta(s)$ to be $\eta^a(s)$ in Theorem 13.2 to conclude (14.12).

Case 2. a is a nonrecurrent trap. Then $\varrho^a(\cdot) \equiv \xi^a(\cdot)$ and by the Corollary to Theorem 12.2, each state i in I^a is either $\prod$-nonrecurrent or $i \not\rightsquigarrow a$. Since a is a trap this means $i \not\rightsquigarrow A$, and so by the discussion following the corollary just cited, each i in I^a which is $\prod$-recurrent must in fact be Φ-recurrent. This has been excluded by Assumption C_0. Hence each i in I^a is $\prod$-nonrecurrent and so

$$r^a = \int_0^\infty \varrho^a(s)\, ds = \int_0^\infty \xi^a(s)\, ds < \infty$$

(see [I; Theorem 6.1], the r^a here being the g^a there). The rest is the same as in Case 1.
Thus in both Case 1 and Case 2 the e^a of the theorem is just the r^a of (14.11).

Case 3. a is a recurrent trap. Then I^a is a $\prod$-recurrent class. By a well-known theorem [1; Theorem II.13.5], there exists $\{e(s),\ s > 0\}$ such that

$$\int_0^\infty e(s)\, ds < \infty, \quad e(s)\textstyle\prod(t) = e(s+t);$$

and consequently $$e(s)\,\Phi(t) \leqslant e(s+t).$$

In fact since all states are stable we may choose any state 0 in I^a and set

$$e_j(s) = {}_0p_{0j}(s), \quad \int_0^\infty e_j(s)\, ds = {}_0p^*_{0j}$$

[1; p. 201]. Hence Theorem 13.2 is applicable with $\Psi = \Phi$ and this choice of $e(s)$, so that e^a is the sequence with

$$e^a_j = {}_0p^*_{0j}, \quad j \in I^a.$$

Remark. It would be interesting to know whether the case of a recurrent trap indeed requires special handling as described above; and if so where is this covered up in the algebraic treatment of other authors.

The next result Theorem 14.4 is one of the two keys to the canonical decomposition. We need an analytical lemma which is essentially known (see [I; § 9] and Neveu [10]);(¹) for a purely analytical proof and general discussion see [II].

LEMMA. *Let σ be finite, nonnegative, nonincreasing in $(0,\infty)$ with $\sigma(0+)\leqslant+\infty$ and $\int_0^1\sigma(s)ds<\infty$. Let $\hat\sigma(\lambda)$ be its Laplace transform:*

$$\hat\sigma(\lambda)=\int_0^\infty e^{-\lambda t}\sigma(t)\,dt,\quad 0<\lambda<\infty.$$

Given also two constants $\delta\geqslant 0$, $p\geqslant 0$. Then there exists a nonnegative measure $E(\cdot)$ on $[0,\infty)$ such that for $0<\lambda<\infty$:

$$[\delta+p\lambda+\lambda\hat\sigma(\lambda)]\,\hat E(\lambda)=1,\tag{14.13}$$

where

$$\hat E(\lambda)=\int_{[0,\infty)}e^{-\lambda t}E(dt).$$

Moreover, E is a finite measure unless $\delta=0$, in which case it is infinite but sigma-finite. If σ is absolutely continuous in (t,∞) for every $t>0$, then E is absolutely continuous except for a point mass at 0 if $\sigma(0)<\infty$.

Proof. By a particular case of P. Lévy's representation of infinitely divisible laws and the associated processes (see Lévy [IV]) there exists an infinitely divisible process $\{Y(v),\,v\geqslant 0\}$ such that if $F(v;\cdot)$ denotes the distribution of $Y(v)$, we have

$$\hat F(v;\lambda)\overset{\text{def}}{=}\int_{[0,\infty)}e^{-\lambda t}F(v;dt)=\mathbf{E}(e^{-\lambda Y(v)})=\exp\left\{-v\left[p\lambda+\int_0^\infty(e^{-\lambda s}-1)\,d\sigma(s)\right]\right\}.$$

Putting

$$\hat u(\lambda)\overset{\text{def}}{=}\lambda\hat\sigma(\lambda)$$

we have

$$e^{-\delta v}\int_{[0,\infty)}e^{-\lambda t}F(v;dt)=e^{-v[\delta+p\lambda+\hat u(\lambda)]}.$$

Integrating this over v in $[0,\infty)$ and setting

$$E(\cdot)\overset{\text{def}}{=}\int_{[0,\infty)}e^{-\delta v}F(v,\cdot)\,dv,\tag{14.14}$$

(¹) Neveu's assertion of a sharper result corresponding to Corollary 1 to Theorem 14.7 below is without substantiation.

we obtain
$$\int_{[0,\infty)} e^{-\lambda t} E(dt) = [\delta + p\lambda + \hat{u}(\lambda)]^{-1}, \tag{14.15}$$

proving (14.13). Letting $\lambda \downarrow 0$ in (14.14) we see that $E([0,\infty)) = \delta^{-1}$ since $\lim_{\lambda \downarrow 0} \hat{u}(\lambda) = 0$, proving that E is an infinite measure if and only if $\delta = 0$, in which case E is still sigma-finite by (14.14). Finally, it is well-known from Lévy's theory that if σ is absolutely continuous in (t, ∞) for every $t > 0$, then $F(v; \cdot)$ is absolutely continuous except for a mass at 0 equal to $\hat{F}(v; +\infty) = e^{-v\sigma(0+)}$ when $p = 0$ and $\sigma(0+) < \infty$, and then $E(\{0\}) = [\delta + \sigma(0+)]^{-1}$ by (14.14). The lemma is completely proved.

Corollary. *Suppose $p = 0$. For (Lebesgue) almost every $t > 0$, we have*
$$\int_{[0,t]} [\delta + \sigma(t-s)]\, E(ds) = 1; \tag{14.16}$$

and for every $t > 0$ the inequality "$\leqslant$" holds above.

This corollary will be sharpened below; see Corollary 1 to Theorem 14.7.

Theorem 14.4. *For each a, $\varrho^a(\cdot)$ and the entrance law $\eta^a(\cdot)$ generated by the e^a in Theorem 14.3 are linked by the following formula, for $t \geqslant 0$:*
$$\varrho^a(t) = \int_{[0,t]} \eta^a(t-s)\, E^a(ds), \tag{14.17}$$

where $E^a(\cdot)$ is a probability measure on $[0,\infty)$, unless a is a recurrent trap in which case it is infinite but sigma-finite. Furthermore $E^a(\cdot)$ is absolutely continuous except for a mass at 0 in case a is nonsticky.

Proof. Case 1. a is not a trap or a is a nonrecurrent trap. Integrating (14.10) over s in $(0,\infty)$ and noting that $e^a = r^a$,
$$e_j^a - \int_0^t \varrho_j^a(s)\, ds = \sum_i e_i^a \left\{ f_{ij}(t) + \int_0^t l_i^a(u)\, \varrho_j^a(t-u)\, du \right\}. \tag{14.18}$$

Using the notation $\langle \cdot, \cdot \rangle$ for the inner product of vectors, we put
$$\sigma^{aa}(t) \overset{\text{def}}{=} \langle e^a, l^a(t) \rangle.$$

Recalling also (13.2) with $e = e^a$ and $\Psi = \Phi$ we may rewrite (14.18) in vector notation as
$$\int_0^t \eta^a(s)\, ds = \int_0^t \varrho^a(u)\, [1 + \sigma^{aa}(t-u)]\, du. \tag{14.19}$$

If we put also $$\theta^{aa}(t) \overset{\text{def}}{=} \langle \eta^a(s), l^a(t-s)\rangle, \quad 0<s<t, \tag{14.20}$$

which is independent of s (see p. 49 of [I]), we have by (14.12)

$$\sigma^{aa}(t) = \int_0^\infty \langle \eta^a(s), l^a(t)\rangle\, ds = \int_0^\infty \theta^{aa}(s+t)\, dt = \int_t^\infty \theta^{aa}(s)\, ds. \tag{14.21}$$

Since $\varrho^a(\cdot)$ is continuous and positive unless a is ephemeral, it follows from (14.19) that $\sigma^{aa}(\cdot)$ is locally integrable. If a is ephemeral then $e^a=0$ so that $\sigma^{aa}(\cdot)\equiv 0$. Hence the Lemma above is applicable to σ^{aa} with $\delta=1$, $p=0$; the corresponding E will be denoted by E^a. It follows from the proof of the lemma that E^a is a probability measure. Taking Laplace transforms in (14.19) we have

$$\hat{\eta}^a(\lambda) = \hat{\varrho}^a(\lambda)[1+\lambda\hat{\sigma}^{aa}(\lambda)],$$

which can be inverted by (14.15) to yield

$$\hat{\varrho}^a(\lambda) = \hat{E}^a(\lambda)\hat{\eta}^a(\lambda).$$

From the uniqueness theorem for Laplace transforms, and the continuity of the functions $\varrho^a(\cdot)$ and $\eta^a(\cdot)$ in $[0,\infty)$, we conclude (14.17) as asserted.

Case 2. a is a recurrent trap.

In this case we have, recalling the handling of this case in Theorem 14.3:

$$e_j^a = \sum_i e_i^a p_{ij}(t) = \sum_i e_i^a \left\{ f_{ij}(t) + \int_0^t \langle l_i^a(u), \varrho_j^a(t-u)\rangle\, du \right\}.$$

Proceeding as before, we obtain

$$\int_0^t \eta^a(s)\, ds = \int_0^t \varrho^a(u)\, \sigma^{aa}(t-u)\, du \tag{14.22}$$

which differs from (14.19) only in that the "1" there is replaced by "0". The Lemma is applicable as before but with $\delta=0$, $p=0$, and the resulting E^a is now an infinite but sigma-finite measure.

We know that E^a is absolutely continuous except for a mass at 0 when $\sigma^{aa}(0)<\infty$; that this last condition is equivalent to a being nonsticky will be shown in Theorem 14.6 below.

In order to combine the two cases above we introduce the symbol

$$\delta^a = \begin{cases} 0, & \text{if } a \text{ is a recurrent trap;} \\ 1, & \text{otherwise.} \end{cases} \tag{14.23}$$

Then we have for each a,

$$[\delta^a + \lambda\hat{\sigma}^{aa}(\lambda)]\,\hat{E}^a(\lambda) = 1;$$

or for almost every t:

$$\int_0^t [1 + \sigma^{aa}(t-s)]\,E^a(ds) = 1. \tag{14.24}$$

COROLLARY. *For each a,*

$$\eta_*^a(\cdot) \overset{\text{def}}{=} \sum_j \eta_j^a(\cdot) = \langle \eta^a(\cdot), 1\rangle$$

is locally integrable.

Proof. We have from (14.19) and (14.22) for every t:

$$\int_0^t \eta_*^a(s)\,ds = \int_0^t \varrho_*^a(s)\,[\delta^a + \sigma^{aa}(t-s)]\,ds \leqslant \int_0^t [1 + \sigma^{aa}(t-s)]\,ds < \infty$$

since σ^{aa} is locally integrable.

THEOREM 14.5. *For each a, $\eta_*^a(\cdot)$ is nonincreasing and*

$$c^a \overset{\text{def}}{=} \eta_*^a(\infty) = \langle \eta^a(t), 1 - L(\infty)\rangle \tag{14.25}$$

for every $t > 0$. This number is also equal to $\varrho_*^a(\infty)$ unless a is a recurrent trap in which case $\eta_*^a(\infty) = 0$, while $\varrho_*^a(\infty) = 1$ for any trap.

Proof. The monotonicity is an immediate consequence of the defining property of an entrance law; and the proof of (14.25) is the same as on p. 50 of [I], although the η there need not be the same as here. Both are properties of any entrance law relative to Φ. Next we have from (14.17):

$$\varrho_*^a(t) = \int_{[0,t]} \eta_*^a(t-s)\,E^a(ds).$$

Letting $t \uparrow \infty$ an recalling that E^a is a probability measure unless a is a recurrent trap, we see that the common value in (14.25) is $\varrho_*^a(\infty)$ except in that case. It is clear from (14.4) that $\varrho_*^a(\infty) = 1$ for any trap a. On the other hand, if a is a recurrent trap, then for every i in I^a, we have $L_i(\infty) = L_i^a(\infty) = 1$ by Theorem 12.3, hence $\eta_*^a(\infty) = 0$ by (14.25).

THEOREM 14.6. *The boundary atom a is sticky if and only if*

$$\sigma^{aa}(0+) = +\infty.$$

Proof. For each $\delta > \varepsilon > 0$, we have clearly

$$\mathbf{P}^a\{S^a \cap (\varepsilon, \delta) \neq \emptyset\} = \langle \varrho^a(\varepsilon), L^a(\delta-\varepsilon)\rangle = \int_{[0,\varepsilon]} \langle \eta^a(\varepsilon - s), L^a(\delta-\varepsilon)\rangle\, E^a(ds)$$

by (14.17). Now by (14.20) and (14.21):

$$\begin{aligned}\langle \eta^a(\varepsilon-s), L^a(\delta-\varepsilon)\rangle &= \int_0^{\delta-\varepsilon} \langle \eta^a(\varepsilon-s), l^a(t)\rangle\, dt \\ &= \int_0^{\delta-\varepsilon} \theta^{aa}(\varepsilon-s+t)\, dt = \sigma^{aa}(\varepsilon-s) - \sigma^{aa}(\delta-s).\end{aligned}$$

Hence $$\mathbf{P}^a\{S^a \cap (0,\delta) \neq \emptyset\} = \lim_{\varepsilon \downarrow 0} \int_{[0,\varepsilon]} [\sigma^{aa}(\varepsilon-s) - \sigma^{aa}(\delta-s)]\, E^a(ds). \tag{14.26}$$

If $\sigma^{aa}(0+) < \infty$, it follows from this that

$$\lim_{\delta \downarrow 0} \mathbf{P}^a\{S^a \cap (0,\delta) \neq \emptyset\} = 0$$

and so a is nonsticky by Definition 12.2. If $\sigma^{aa}(0+) = +\infty$, then $E^a(\{0\}) = 0$ by the lemma, it follows from (14.26) and (14.16) that

$$\mathbf{P}^a\{S^a \cap (0,\delta) \neq \emptyset\} = \lim_{\varepsilon \downarrow 0} \int_{[0,\varepsilon]} \sigma^{aa}(\varepsilon - s)\, E^a(ds) = 1.$$

This being true for every $\delta > 0$, we have (12.5) and so a is sticky. Q.e.d.

Generalizing the definitions in (14.20) and (14.21), we put for $a \in \mathbf{A}$, $b \in \mathbf{A}$:

$$\theta^{ab}(t) \overset{\text{def}}{=} \langle \eta^a(s), l^b(t-s)\rangle, \quad 0 < s < t; \qquad \sigma^{ab}(t) \overset{\text{def}}{=} \int_t^\infty \theta^{ab}(s)\, ds. \tag{14.27}$$

It follows by (14.12) that

$$\sigma^{ab}(t) = \int_0^\infty \theta^{ab}(s+t)\, ds = \begin{cases} \displaystyle\int_0^\infty \langle \eta^a(s), l^b(t)\rangle\, dt = \langle e^a, l^b(t)\rangle, \\ \displaystyle\int_0^\infty \langle \eta^a(t), l^b(s)\rangle\, ds = \langle \eta^a(t), L^b(\infty)\rangle. \end{cases} \tag{14.28}$$

Furthermore, for $t > 0$:

$$\eta^a_*(t) = \langle \eta^a(t), 1 - L(\infty) + \sum_{b\in\mathbf{A}} L^b(\infty)\rangle = c^a + \sum_{b\in\mathbf{A}} \sigma^{ab}(t). \tag{14.29}$$

From here on, as a convention in notation, a Lebesgue-Stieltjes integral such as $\int_0^t \dots dE(s)$ shall mean $\int_{[0,t]} \dots E(ds)$ for finite t and $\int_{[0,\infty)} \dots E(ds)$ for $t = +\infty$. Furthermore, $E(0)$ will be written for $E(\{0\})$.

We proceed to derive new basic relations for ϱ^a and F^{ab}, $a \in \mathbf{A}$, $b \in \mathbf{A} - \{a\}$. Note that some results may be vacuously true for an ephemeral a or a trap. We begin with

$$\mathbf{P}_i\{\beta^a \leqslant t, x(\beta^a) = b\} = L_i^b(t) + \int_0^t F^{ab}(t-s)\, dL_i^a(s). \tag{14.30}$$

We observe that the probability on the left side above is equal to

$$\mathbf{P}_i\{\alpha = \beta^a \leqslant t;\, x(\beta^a) = b\} + \mathbf{P}_i\{\alpha < \beta^a \leqslant t;\, x(\beta^a) = b\}.$$

The first term is $L_i^b(t)$ by definition, and the second is by the Strong Markov property equal to

$$\mathbf{P}_i\{\alpha < t;\, x(\alpha) = a;\, \beta^a \leqslant t;\, x(\beta^a) = b\} = \int_0^t \mathbf{P}^a\{\beta^a \leqslant t-u;\, x(\beta^a) = b\}\, d\mathbf{P}_i\{\alpha \leqslant u\}$$
$$= \int_0^t F^{ab}(t-u)\, dL_i^a(u).$$

Hence (14.30) is proved.

Next, we have

$$\mathbf{P}^a\{s < \beta^a \leqslant s+t;\, x(\beta^a) = b\} = \sum_i \mathbf{P}^a\{s < \beta^a;\, x(s) = i\}\, \mathbf{P}_i\{\beta^a \leqslant t;\, x(\beta^a) = b\}.$$

Using Definition 14.2 and (14.30), this is

$$F^{ab}(s+t) - F^{ab}(s) = \langle \varrho^a(s), L^b(t)\rangle + \int_0^t \langle \varrho^a(s), l^a(u)\rangle\, F^{ab}(t-u)\, du. \tag{14.31}$$

Let us introduce the further notation

$$\varrho^{ab}(t) \stackrel{\text{def}}{=} \langle \varrho^a(t), L^b(\infty)\rangle. \tag{14.32}$$

Since each $\varrho^a(\cdot)$ is continuous and the right side above is dominated by $\langle \xi^a(t), 1\rangle$ which converges uniformly in every finite interval it follows that $\varrho^{ab}(\cdot)$ is continuous.

THEOREM 14.7. *We have, for every $t > 0$:*

$$E^a(t) = 1 - \varrho^{aa}(t); \tag{14.33}$$

$$E^a(t)\, F^{ab}(\infty) = \varrho^{ab}(t) + F^{ab}(t); \tag{14.34}$$

$$F^{ab}(t) = \int_0^t [F^{ab}(\infty) - \sigma^{ab}(t-s)]\, dE^a(s). \tag{14.35}$$

Proof. From (14.17) and (14.32) we have

$$\varrho^{ab}(t) = \int_0^t \sigma^{ab}(t-s)\, dE^a(s). \tag{14.36}$$

In particular, by (14.24), and the uniqueness of Laplace transforms:

$$\varrho^{aa}(t) = \int_0^t \sigma^{aa}(t-s)\, dE^a(s) = 1 - \delta^a E^a(t)$$

first for (Lebesgue) almost every t but then for every $t>0$ since both extreme terms above are continuous in $(0,\infty)$. Formula (14.33) is thus a sharp form of (14.24). Now letting $t \uparrow \infty$ and then replacing s by t in (14.31), we obtain

$$F^{ab}(\infty) - F^{ab}(t) = \varrho^{ab}(t) + \varrho^{aa}(t)\, F^{ab}(\infty).$$

This is (14.34) on account of (14.33). It is also (14.35) on account of (14.36).

COROLLARY 1. $\forall t>0$: $\int_0^t [\delta^a + \sigma^{aa}(t-s)]\, dE^a(s) = 1$.

COROLLARY 2. *$F^{ab}(\cdot)$ is absolutely continuous.*

Equation (14.34) becomes perhaps more interesting if it is divided through by $F^{ab}(\infty)$, supposed to be positive; showing then that the resulting right side does not depend on b and defines the basic measure $E^a(\cdot)$. A similar relation involving an arbitrary j may be recorded as follows:

$$E^a(t) = \frac{\int_0^t \varrho_j^a(s)\, ds + \sum_i \varrho_i^a(t) \int_0^\infty f_{ij}(s)\, ds}{\int_0^\infty \varrho_j^a(s)\, ds} = \frac{\mathbf{E}^a\{\mu[S_j \cap (0, \alpha_t \wedge \beta^a)]\}}{\mathbf{E}^a\{\mu[S_j \cap (0, \beta^a)]\}}.$$

This is proved by integrating the equation

$$\varrho^a(s+t) = \varrho^a(s)\,\Phi(t) + \int_0^t \langle \varrho^a(s), l^a(u)\rangle\, \varrho^a(t-u)\, du$$

over t. It would be interesting to understand the meaning of this "equilibrium property" of $E^a(\cdot)$ with respect to j in $\mathbf{I}$ as well as to b in $\mathbf{A}$.

An interesting consequence of (14.35) is its limit as $t \downarrow 0$:

$$F^{ab}(0) = E^a(0)\,[F^{ab}(\infty) - \sigma^{ab}(0)]. \tag{14.37}$$

THEOREM 14.8. *For any $a \neq b$, we have*

$$\sigma^{ab}(0) \leqslant F^{ab}(\infty) \leqslant 1. \tag{14.38}$$

Proof. Integrating (14.31) over s in $(0,\infty)$, we obtain

$$F^{ab}(\infty)\,t-\int_0^t F^{ab}(s)\,ds=\langle e^a, L^b(t)\rangle+\int_0^t \sigma^{aa}(u)\,F^{ab}(t-u)\,du,$$

or using (14.28) to transform the first term on the right side,

$$F^{ab}(\infty)\,t=\int_0^t \{\sigma^{ab}(u)+[1+\sigma^{aa}(u)]\,F^{ab}(t-u)\}\,du.$$

Dividing through by t and letting $t\downarrow 0$, observing that $\sigma^{ab}(\cdot)$ is nonincreasing we infer (14.38).

COROLLARY.
$$\lim_{t\downarrow 0}\,[\eta^a_*(t)-\sigma^{aa}(t)]<\infty. \tag{14.39}$$

Combining Theorems 14.6 and 8 we conclude that the matrix $(\sigma^{ab}(0))$, $(a,b)\in \mathbf{A}\times\mathbf{A}$, has finite elements off the diagonal while a diagonal element is finite or infinite according as it corresponds to a nonsticky or sticky atom. This simple result used to be an obscure point in previous investigations ([7], [10], [VIII]).

§ 15. Canonical decomposition

As explained above the quantities ϱ^a and F^{ab} serve the purpose of separating the banners from each other as long as possible; now it is necessary to link them together. The leading formula, which is the second key to the canonical decomposition, the first being Theorem 14.4, is given below.

THEOREM 15.1. *For each a, we have*

$$\xi^a_j(t)=\varrho^a_j(t)+\sum_{b\neq a}\int_0^t \xi^b_j(t-s)\,dF^{ab}(s),\quad j\in\mathbf{I}^a. \tag{15.1}$$

Remark. This equation need not be valid for an arbitrary j; for example if $j\in\mathbf{I}^b\backslash\mathbf{I}^a$ and $a\rightsquigarrow b$; then the left member is 0 but the right member is positive.

Proof. The meaning of (15.1) is obious: we rewrite it in terms of random variables:

$$\mathbf{P}^a\{x(t)=j\}=\mathbf{P}^a\{\beta^a>t;x(t)=j\}+\sum_{b\neq a}\mathbf{P}^a\{\beta^a\leqslant t;\beta^a\in\widetilde{S^b};x(t)=j\}.$$

Let us observe at once the notational quirk which necessitates the use of "$\beta^a\in\widetilde{S^b}$", or more clearly perhaps: "$\forall\delta>0$: $S^b\cap(\beta^a,\beta^a+\delta)\neq\emptyset$" instead of the obvious "$x(\beta^a)=b$" because the latter would have the wrong meaning when $\beta^a=0$. The rigorous proof of the preceding

equation is of course, as always, done by invoking the Strong Markov property, but since Theorems 4.3 and 4.4 of [I] were formulated a little too narrowly for the present purpose we shall indicate the minor modifications needed.

1°. The cited theorems were stated for τ^a which is the first time the boundary is reached if it is reached at the atom a, otherwise equal to $+\infty$. However, the same results hold if this is replaced by any optional random variable τ such that $x(\tau)=a$ where $\tau<+\infty$. By definition then there exist $\tau_0<\tau$ and $\tau_n \uparrow \tau$ such that the jump chain $\chi_n=x(\tau_n)$ starting at time $\tau_0(=\tau_0(\omega))$ reaches a at time τ. The proofs given for the cited theorems are valid if this $\{\chi_n\}$ is used there instead of the jump chain starting at time 0.

2°. The cited theorem did not deal with the situation where the given $\{x_t\}$ starts at a boundary atom and is an open Markov chain. However, nothing in the proofs, as revised in 1° above, is changed on the set $\{\tau>0\}$ (in which case the chain starts *in effect* at an ordinary state as far as τ is concerned).

3°. Thus we are left with that case $\{\tau=0\}$ under the initial condition say $x(0)=a$. This is settled by the following simple lemma.

LEMMA 1. *For each Λ in $\mathfrak{F}^0$ (the Borel field generated by $\{x_t\}$) we have*

$$\mathbf{P}\{0\in\overline{S^b};\Lambda\}=F^{ab}(0)\,\mathbf{P}^b(\Lambda).$$

Proof. The meaning of this is again clear: on the set $\{0\in\overline{S^b}\}$ the process acts as if it started at b (rather than a). To prove it we define for each positive integer m:

$$\tau_m(\omega)=\inf\{t:t>m^{-1};x(t,\omega)=b\};$$

then $\tau_m\downarrow 0$ on $\{0\in\overline{S^b}\}$. If $0<\varepsilon<t$. then since $\{0\in\overline{S^b}\}\in\mathfrak{F}_{\tau_m}$ for each m, we have by the Strong Markov property:

$$\mathbf{P}^a\{0\in\overline{S^b};\tau_m\leqslant\varepsilon;x(t)=j\}=\int_0^\varepsilon \xi_j^b(t-s)\,d\mathbf{P}^a\{0\in\overline{S^b};\tau_m\leqslant s\}.$$

Letting $m\uparrow\infty$ we obtain

$$\mathbf{P}^a\{0\in\overline{S^b};x(t)=j\}=\mathbf{P}^a\{0\in\overline{S^b}\}\,\xi_j^b(t)=F^{ab}(0)\,\xi_j^b(t).$$

This being true for every $t>0$, and $\{0\in\overline{S^b}\}\in\wedge_{m=1}^{\infty}\mathfrak{F}_{m^{-1}}$, the lemma follows.

Theorem 15.1 is now proved by applying the amendments 1°, 2°, 3° to β^a. Thus, the typical term in the sum at the beginning of the proof is further split into

$$\begin{aligned}\mathbf{P}^a\{\beta^a=0\in\overline{S^b};x(t)=j\}&+\mathbf{P}^a\{0<\beta^a\leqslant t;x(\beta^a)=b;x(t)=j\}\\&=F^{ab}(0)\,\xi_j^b(t)+\int_{(0,t]}\xi_j^b(t-s)\,F^{ab}(ds).\end{aligned}$$

Q.e.d.

Taking Laplace transforms in (15.1), we obtain

$$\hat{\xi}^a(\lambda) = \hat{\varrho}^a(\lambda) + \sum_{b \neq a} \hat{F}^{ab}(\lambda)\, \hat{\xi}^b(\lambda), \tag{15.2}$$

where
$$\hat{\xi}^a(\lambda) = \int_0^\infty e^{-\lambda t} \xi^a(t)\, dt, \qquad \hat{\varrho}^a(\lambda) = \int_0^\infty e^{-\lambda t} \varrho^a(t)\, dt,$$

$$\hat{F}^{ab}(\lambda) = \int_0^\infty e^{-\lambda t}\, dF^{ab}(t).$$

Putting $\hat{F}^{aa}(\lambda) \equiv 0$ for each $a \in \mathbf{A}$, and introducing the matrix

$$\hat{F}(\lambda) \stackrel{\text{def}}{=} (\hat{F}^{ab}(\lambda)), \quad (a, b) \in \mathbf{A} \times \mathbf{A}$$

as well as the vectors $\hat{\xi}(\lambda)$, etc. where $\hat{\xi}(\lambda) = \{\hat{\xi}^a(\lambda), a \in \mathbf{A}\}$, we may write (15.2) in matrix notation as follows:

$$[I - \hat{F}(\lambda)]\, \hat{\xi}(\lambda) = \hat{\varrho}(\lambda). \tag{15.3}$$

Note that the vectors as well as matrices above are indexed on the finite set $\mathbf{A}$ of superscripts, the subscripts j in $\mathbf{I}^a$ being understood.

The next task is that of solving for $\hat{\xi}$ from (15.3) and that is done by the lemma below. Although it is but a special case of a "recurring theorem" (see Taussky [VI]) in its definitive form, we shall spell out a constructive proof using the theory of Markov chains.

LEMMA 2. *Let $P = (p_{ab})$, $(a, b) \in \mathbf{A} \times \mathbf{A}$, where $\mathbf{A}$ is a finite set, be a substochastic matrix. A necessary and sufficient condition that $I - P$ be invertible is: there does not exist a subset $\mathbf{C}$ of $\mathbf{A}$ such that $P|_\mathbf{C}$ (= the restriction of P to $\mathbf{C} \times \mathbf{C}$) is stochastic. The inverse has nonnegative elements when it exists.*

Proof. Consider the stochastic completion $\tilde{P}$ of P by ϑ (see [I; pp. 22–23]). Unless P is stochastic, $\tilde{P}$ is on the enlarged index set $\tilde{\mathbf{A}} = \mathbf{A} \cup \{\vartheta\}$. A discrete parameter Markov chain with minimial state space $\tilde{\mathbf{A}}$ and one-step transition matrix $\tilde{\mathbf{P}}$ will have ϑ as an absorbing state, and any state a such that $a \rightsquigarrow \vartheta$ will be inessential, hence nonrecurrent.

Case 1. $\tilde{P} \neq P$. Suppose the condition of the lemma holds, then every state except ϑ is nonrecurrent. For if there were any recurrent state distinct from ϑ, there would be a recurrent class $\mathbf{C}$ such that $\vartheta \notin \mathbf{C}$, thus $P|_\mathbf{C} = \tilde{P}|_\mathbf{C}$ would be stochastic, contrary to hypothesis. Hence for any a and b in $\mathbf{A}$ we have, in familiar notation:

$$s_{ab} \overset{\text{def}}{=} \sum_{n=0}^{\infty} \tilde{p}_{ab}^{(n)} < \infty.$$

Since $\tilde{p}_{ab}^{(n)} = p_{ab}^{(n)}$ for every a and b in $\mathbf{A}$, we have

$$s_{ab} = \sum_{n=0}^{\infty} p_{ab}^{(n)} < \infty. \tag{15.4}$$

Let S denote the matrix (s_{ab}), $(a, b) \in \mathbf{A} \times \mathbf{A}$. It is easy to verify that

$$(I-P)S = I = S(I-P), \tag{15.5}$$

the second equation being of course also a consequence of the first. Hence $I-P$ has the inverse $S \geqslant 0$.

Case 2. $\tilde{P} = P$. Then under the condition of the lemma every state in $\mathbf{A}$ is nonrecurrent with respect to P by the same argument as before, and so (15.4) is still true. The rest is the same as before.

The sufficiency of the condition is proved.

Now suppose the condition of the lemma is not fulfilled, namely that there exists a subset $\mathbf{C}$ of $\mathbf{A}$ such that $P|_{\mathbf{C}}$ is stochastic. Then define the vector $w = \{w_a, a \in \mathbf{A}\}$ as follows: $w_a = 1$ or 0 according as $a \in \mathbf{C}$ or $a \in \mathbf{A} - \mathbf{C}$. Clearly we have $w \neq 0$ and $(I-P)w = 0$ so that $I-P$ is not invertible.

Remark. The "sufficiency" part of the lemma and its proof above can be extended to an infinite index set $\mathbf{A}$, but the resulting "inverse" in the sense of (15.5) is not a true inverse operator because the usual multiplication of infinite matrices, even when it is defined, is not necessarily associative. Thus $(I-P)x = y$ is not equivalent to or even implies $x = (I-P)^{-1}y$. An appropriate extension may be needed for the boundary theory with infinitely many atoms.

We are now in position to formulate the theorem on canonical decomposition.

THEOREM 15.2. *If all boundary atoms are distinguishable then for every* λ: $0 < \lambda < \infty$, *the matrix* $I - \hat{F}(\lambda)$ *in* (15.3) *has a nonnegative inverse so that we have*

$$\hat{\xi}(\lambda) = [I - \hat{F}(\lambda)]^{-1} \hat{\varrho}(\lambda) = [I - \hat{F}(\lambda)]^{-1} \hat{E}(\lambda)\, \hat{\eta}(\lambda), \tag{15.6}$$

where $\hat{E}(\lambda)$ *is the diagonal matrix with entries* $\{\hat{E}^a(\lambda), a \in \mathbf{A}\}$,

$$\hat{E}^a(\lambda) = \int_0^\infty e^{-\lambda t}\, dE^a(t), \qquad \hat{\eta}^a(\lambda) = \int_0^\infty e^{-\lambda t}\, \eta^a(t)\, dt,$$

and E^a *and* η^a *are given in Theorem 14.4.*

Proof. Fix a λ and apply Lemma 2 with $P=\hat{F}(\lambda)$. Suppose that there exists $\mathbf{C}\subset\mathbf{A}$ such that $\hat{F}(\lambda)|_{\mathbf{C}}$ is stochastic then there exists $\mathbf{C}_0\subset\mathbf{C}$ such that $\hat{F}(\lambda)|_{\mathbf{C}_0}$ is the one-step transition matrix of a discrete parameter Markov chain where state space $\mathbf{C}_0$ is a recurrent class. Since $\hat{F}(\lambda)$ has zero elements on the diagonal, $\mathbf{C}_0$ must contain at least two states, thus

$$\sum_{b\in\mathbf{C}_0-\{a\}}\hat{F}^{ab}(\lambda)=1,\quad a\in\mathbf{C}_0. \tag{15.7}$$

Furthermore, since $\lambda>0$, $\hat{F}^{ab}(\lambda)\leqslant F^{ab}(\infty)$ and $\sum_{b\neq a}F^{ab}(\infty)\leqslant 1$, (15.7) is possible for each a if and only if

$$\hat{F}^{ab}(\lambda)=F^{ab}(0)=F^{ab}(\infty)\quad\text{and}\quad\sum_{b\neq a}F^{ab}(0)=1.$$

It follows then by (14.4) that $\varrho^a(\cdot)\equiv 0$ and consequently (15.1) is reduced to

$$\xi_j^a(t)=\sum_{b\in\mathbf{C}_0-\{a\}}F^{ab}(0)\,\xi_j^b(t),\quad a\in\mathbf{C}_0,$$

for each $t>0$ and $j\in\mathbf{I}^a$. This means the matrix $(F^{ab}(0))$, $(a,b)\in\mathbf{C}_0\times\mathbf{C}_0$, with $F^{aa}(0)=0$ for each a, is the one-step transition matrix of a discrete parameter recurrent Markow chain and that $\{\xi_j^a(t), a\in\mathbf{C}_0\}$ for fixed t and j is a harmonic (regular) function on $\mathbf{C}_0$ relative to this matrix. Such a function must be a constant, namely:

$$\xi_j^a(t)=\xi_j^b(t),\quad j\in\mathbf{I}^a. \tag{15.8}$$

It follows that

$$1=\sum_{j\in\mathbf{I}^a}\xi_j^a(t)=\sum_{j\in\mathbf{I}^a}\xi_j^b(t).$$

Hence $\mathbf{I}^b\subset\mathbf{I}^a$ and so $\mathbf{I}^a=\mathbf{I}^b$ for every a and b in $\mathbf{C}_0$, since a and b are interchangeable above. Thus (15.8) is true for every j in $\mathbf{I}^a=\mathbf{I}^b$ and every t, and so a and b are indistinguishable by definition (see [I; p. 44]), contrary to hypothesis.

We have thus proved that the condition of the lemma is fulfilled for $\hat{F}(\lambda)$ for each $\lambda>0$, and therefore $I-\hat{F}(\lambda)$ is invertible. This proves the first equation in (15.6); the second follows from (14.17). Theorem 15.2 is completely proved.

It is instructive to compare this proof with that of Theorem 5.5 of [I] to see where progress has been made.

Before proceeding further we will stop for a minor generalization in order to include the case where $\prod$ is substochastic (see pp. 23–24, 25–26 of [I]).[1] Instead of Assumption C_0 in § 14 we make the following slightly weaker one.

ASSUMPTION C_1. *There is at most one Φ-recurrent state (and this is then necessarily Φ-absorbing).*

[1] It is not clear why the substochastic case causes trouble in Williams [VIII].

This state is to be denoted by θ and not included in $\mathbf{I}$, and $\mathbf{I}_\theta = \mathbf{I} \cup \{\theta\}$ is to be the state space if θ is present under C_1. We have [I; (2.7)]

$$f_{i\theta} \equiv 0, \quad f_{\theta i} \equiv 0, \quad f_{\theta\theta} \equiv 1, \quad L_\theta \equiv 0. \tag{15.9}$$

It follows that θ does not belong to $\mathbf{I} - \bigcup_{a \in \mathbf{A}} \mathbf{I}^a$ provided that $\mathbf{P}\{x(0) = \theta\} = 0$; in other words, with probability one θ does not appear before the boundary is reached. It may belong to some $\mathbf{I}^a$ and not to others. If it belongs to $\mathbf{I}^a$, then a cannot be recurrent. For such an θ, we have by (14.10):

$$\varrho_\theta^a(s+t) = \varrho_\theta^a(s) + \sum_{i \in \mathbf{I}_\theta} \varrho_i^a(s) f_{i\theta}(t) + \int_0^t \left[\sum_{i \in \mathbf{I}_\theta} \varrho_i^a(s) l_i^a(u) \right] \varrho_\theta^a(t-u)\, du.$$

It is seen at once that the first sum above is equal to zero, and in the second the summation may be replaced by $i \in \mathbf{I}$, on account of (15.9). Since a is not recurrent $r_i^a = \int_0^\infty \varrho_i^a(t)\, dt < \infty$ for every $i \in \mathbf{I}^a - \{\theta\}$, it follows by integrating the equation above that

$$\int_0^\infty [\varrho_\theta^a(s+t) - \varrho_\theta^a)]\, ds = \int_0^t \sigma^{aa}(u)\, \varrho_\theta^a(t-u)\, du.$$

Since $\varrho_\theta^a(\cdot) \nearrow$, the limit
$$d^a \stackrel{\text{def}}{=} \lim_{t \uparrow \infty} \varrho_\theta^a(t) \tag{15.10}$$

exists and the equation above reduces to

$$d^a t = \int_0^t \varrho_\theta^a(u)\, du\, [1 + \sigma^{aa}(t-u)]\, du. \tag{15.11}$$

Comparing this with (14.19), we see that we should set

$$\eta_\theta^a(\cdot) \stackrel{\text{def}}{=} d^a \tag{15.12}$$

in order that (14.19) may be valid for θ as well as the other states in $\mathbf{I}^a$. With this definition (14.17) is valid as follows:

$$\varrho_\theta^a(t) = \int_0^t \eta_\theta^a(t-s)\, E^a(ds) = d^a E^a(t). \tag{15.13}$$

Now (15.1), (15.2) and (15.3) are valid for the subscript θ as well as the others in $\mathbf{I}^a$. Finally (14.29) becomes

$$\sum_{i \in \mathbf{I}_\theta} \eta_i^a(t) = d^a + c^a + \sum_{b \in \mathbf{A}} \sigma^{ab}(t).$$

If a is recurrent then $\theta \notin \mathbf{I}^a$ and we may set $d^a = 0$.

Briefly, the above discussion goes to show that if there is only one Φ-recurrent state, or more generally if all Φ-recurrent states are merged into one absorbing state, it acts like an atom in the recurrent part of the boundary as to be expected from the Martin theory.

Assumption C_1 is justified as follows. A Φ-recurrent state belongs to the set Z of states from which with probability one the boundary will not be reached at all [I; Theorem 3.2], so that after entering such a state the process is controlled by Φ above. Such states form a stochastically closed set which splits into possibly infinitely many disjoint Φ-recurrent classes. Each class may be treated as the θ above as an atom on the recurrent part of the boundary; or in the informal language used before, as a banner trap under which the states of that class line up according to the law of transition Φ. By merging all these classes into one single trap θ in Assumption C_1, we are just stopping the process at the recurrent part of the boundary—not so much because it can be totally ignored but because its behavior from there on is well known and may be separated from the rest of the study in the name of convenience.

§ 16. Construction

We shall lead up to the so-called "construction theorem" by reviewing the components and steps, suitably algebraicized, which enter into the canonical decomposition (15.6).

Let Q be given satisfying Assumption A and let $\Phi(\cdot)$ be the minimal solution associated with it. Put

$$z \overset{\text{def}}{=} L(\infty) = \lim_{t\uparrow\infty}[I-\Phi(t)]\,1\,\lim_{\lambda\downarrow 0}[I-\lambda\hat{\Phi}(\lambda)]\,1; \tag{16.1}$$

z is an exit sequence relative to Φ which is maximal in the sense of Theorem 13.3. Let $\mathbf{A}$ be a finite index set; for each a in $\mathbf{A}$ let z^a be an exit sequence relative to Φ such that

$$z = \sum_{a\in\mathbf{A}} z^a. \tag{16.2}$$

Let
$$z^a(\lambda) \overset{\text{def}}{=} [I-\lambda\hat{\Phi}(\lambda)]\,z^a.$$

It follows from Theorem 13.3 that

$$\lim_{\lambda\downarrow 0}\lambda\hat{\Phi}(\lambda)\,z^a = 0. \tag{16.3}$$

For each a in $\mathbf{A}$ let e^a be an entrance sequence relative to Φ. This means in terms of Laplace transforms:

$$\forall\lambda\geqslant 0\colon\ e^a \geqslant e^a\lambda\hat{\Phi}(\lambda).$$

Assume that
$$\lim_{\lambda \downarrow 0} e^a \lambda \hat{\Phi}(\lambda) = 0 \tag{16.4}$$

and define the entrance law η^a as follows:

$$\hat{\eta}^a(\lambda) \overset{\text{def}}{=} k^a e^a [I - \lambda \hat{\Phi}(\lambda)], \tag{16.5}$$

where k^a is a positive constant to be fixed later. We set also

$$c^a k^a c_0^a \overset{\text{def}}{=} \langle \lambda \hat{\eta}^a(\lambda), 1 - z \rangle \tag{16.6}$$

which does not depend on λ; and

$$\hat{\sigma}^{ab}(\lambda) \overset{\text{def}}{=} \langle \hat{\eta}^a(\lambda), z^b \rangle, \qquad \hat{u}^{ab}(\lambda) \overset{\text{def}}{=} \lambda \hat{\sigma}^{ab}(\lambda). \tag{16.7}$$

Now we assume that, for each a and $0 \leqslant \lambda < \infty$:

$$\hat{u}^{aa}(\lambda) < \infty; \tag{16.8}$$

and for $a \neq b$:
$$\hat{u}^{ab}(\infty) \overset{\text{def}}{=} \lim_{\lambda \uparrow \infty} \hat{u}^{ab}(\lambda) < \infty. \tag{16.9}$$

For each pair a and b, $a \neq b$, we choose a constant Ω_0^{ab} such that

$$\hat{u}^{ab}(\infty) \leqslant k^a \Omega_0^{ab} < \infty. \tag{16.10}$$

Let $\Omega_0^{aa} = 0$ for each a, and

$$\mathbf{A}_0 \overset{\text{def}}{=} \{a \in \mathbf{A} : c_0^a + \sum_b \Omega_0^{ab} = 0\}$$

and
$$\delta^a \overset{\text{def}}{=} \begin{cases} 0 & \text{if} \quad a \in \mathbf{A}_0, \\ 1 & \text{if} \quad a \in \mathbf{A} - \mathbf{A}_0. \end{cases}$$

Now fix the constant k^a for each a so that

$$k^a c_0^a + k^a \sum_b \Omega_0^{ab} = \delta^a. \tag{16.11}$$

Write Ω^{ab} for $k^a \Omega_0^{ab}$ and let Ω be the matrix (Ω^{ab}), $(a, b) \in \mathbf{A} \times \mathbf{A}$. Furthermore, put

$$\hat{E}^a(\lambda) \overset{\text{def}}{=} \frac{1}{\delta^a + \hat{u}^{aa}(\lambda)} \tag{16.12}$$

and
$$\hat{\varrho}^a(\lambda) \overset{\text{def}}{=} \hat{E}^a(\lambda)\, \hat{\eta}^a(\lambda); \tag{16.13}$$

$$\hat{F}^{aa}(\lambda) \overset{\text{def}}{=} 0; \qquad \hat{F}^{ab}(\lambda) \overset{\text{def}}{=} \hat{E}^a(\lambda)\, (\Omega^{ab} - \hat{u}^{ab}(\lambda)), \quad a \neq b. \tag{16.14}$$

Let $\hat{F}(\lambda)$ be the matrix $(F^{ab}(\lambda))$. If there exists a subset $\mathbf{C}$ of $\mathbf{A}$ such that $(\hat{F}(0))|_{\mathbf{C}}$ is stochastic, then we decree that all indices in $\mathbf{C}$ be identified. When this has been done $I-\hat{F}(\lambda)$ will be invertible by Lemma 2 of § 15. Let $\hat{D}(\lambda)$ be the diagonal matrix with entries $(\hat{u}^{aa}(\lambda), a\in\mathbf{A})$ and let $\tilde{I}$ be the diagonal matrix with entries $(\delta^a, a\in\mathbf{A})$. Finally, we put

$$\begin{aligned}\hat{\xi}(\lambda) &\overset{\text{def}}{=} [I-\hat{F}(\lambda)]^{-1}\hat{\varrho}(\lambda) = [I-\hat{F}(\lambda)]^{-1}\hat{E}(\lambda)\hat{\eta}(\lambda)\\ &= [I-\hat{F}(\lambda)]^{-1}[\tilde{I}+\hat{D}(\lambda)]^{-1}\hat{\eta}(\lambda) = \hat{M}(\lambda)\hat{\eta}(\lambda),\end{aligned} \tag{16.15}$$

where $\hat{M}(\lambda)=(\hat{M}^{ab}(\lambda))$, $(a,b)\in\mathbf{A}\times\mathbf{A}$ is the matrix defined below:

$$\hat{M}(\lambda) \overset{\text{def}}{=} [I-\hat{F}(\lambda)]^{-1}[\tilde{I}+\hat{D}(\lambda)]^{-1}; \tag{16.16}$$

and
$$\hat{\prod}(\lambda) \overset{\text{def}}{=} \hat{\Phi}(\lambda) + \sum_{a\in\mathbf{A}} \hat{z}^a(\lambda)\hat{\xi}^a(\lambda) = \hat{\Phi}(\lambda) + \sum_{a\in\mathbf{A}}\sum_{b\in\mathbf{A}} \hat{z}^a(\lambda)\hat{M}^{ab}(\lambda)\hat{\eta}^b(\lambda). \tag{16.17}$$

THEOREM 16.1. *The $\hat{\prod}(\lambda)$ so constructed is the Laplace transform (resolvent) of a stochastic transition matrix function $\prod(t)$:*

$$\hat{\prod}(\lambda) = \int_0^\infty e^{-\lambda t}\prod(t)\,dt.$$

Conversely, under Assumptions A, B′, C_0 *and* D *every such $\prod$ may be constructed in the manner described above.*

Proof. The "converse" part, somewhat vaguely stated here, is just an algebraic re-statement of Theorem 15.2 and Theorem 5.1 of [I] with $z^a = L^a(\infty)$.

The proof of the "direct" part will be sketched. It is merely a matter of algebraic verification based on the resolvent equations, and is quite similar to pp. 67–68 of [I]. Dropping the "^" on Laplace transforms, we need to verify (8.26) of [I], i.e.

$$(\lambda-\mu)M(\lambda)\Theta(\lambda,\mu)M(\mu) = M(\mu)-M(\lambda)$$

or equivalently, since, $U(\lambda)$ being the matrix $(u^{ab}(\lambda))$,

$$(\lambda-\mu)\Theta(\lambda,\mu) = \Theta(\mu)-\Theta(\lambda) = U(\lambda)-U(\mu)$$

(see (8.25) of [I] and the equations preceding it), that

$$M(\lambda)[U(\lambda)-U(\mu)]M(\mu) = M(\mu)-M(\lambda). \tag{16.18}$$

Since $M(\lambda)^{-1}$ exists by (16.16), this in turn is equivalent to

$$U(\lambda)-U(\mu)=M(\lambda)^{-1}[M(\mu)-M(\lambda)]M(\mu)^{-1}=M(\lambda)^{-1}-M(\mu)^{-1}. \tag{16.19}$$

Now by simple inspection, one sees that

$$M(\lambda)^{-1}=[\tilde{I}+D(\lambda)][I-F(\lambda)]=\tilde{I}-\Omega+U(\lambda). \tag{16.20}$$

For, on the diagonal, $F(\lambda)$ and Ω are both zero while $D(\lambda)$ is by definition the diagonal part of $U(\lambda)$; off the diagonal, the second equation in (16.20) reduces to

$$[\tilde{I}+D(\lambda)]F(\lambda)=\Omega-U(\lambda)$$

which is just (16.14). Since $\tilde{I}-Q$ does not depend on λ, (16.19) follows from (16.20) and so the resolvent equation for $\prod(\lambda)$ is verified (see the calculation after (8.26) in [I]).

To verify "stochasticity" (the "norm condition"), we write

$$I=M(\lambda)M(\lambda)^{-1}=M(\lambda)[\tilde{I}-\Omega+U(\lambda)];$$

hence by (16.11) and (14.29) (Laplace-transformed):

$$\begin{aligned}1=I1&=M(\lambda)[\tilde{I}1-\Omega 1+U(\lambda)1]\\&=M(\lambda)[c+U(\lambda)1]=M(\lambda)\langle\lambda\eta(\lambda),1\rangle=\langle\lambda\xi(\lambda),1\rangle.\end{aligned} \tag{16.21}$$

This is necessary and sufficient for $\prod(\cdot)$ to be stochastic. Theorem 16.1 is completely proved.

COROLLARY. *To construct a substochastic $\prod(\cdot)$, choose any $d_0^a\geqslant 0$ for each a and choose k^a so that instead of* (16.11), *we have*

$$k^a(d_0^a+c_0^a+\sum_b\Omega_0^{ab})=\delta^a.$$

Apart from this no change is needed in the procedure. Conversely every substochastic $\prod(\cdot)$ can be constructed in this manner; and it will be strictly substochastic if and only if $d^a>0$ for some a.

The idea of the above extension to the substochstic case is of course a trivial and familiar one: one first constructs the stochastic completion by adjoining a new index to **I**, and the restriction of the completion to **I** will be the most general substochastic case.

There is another complement, the so-called "extension to the boundary" discussed by previous authors. The idea is to construct a Markov chain x^* on the enlarged state space $\mathbf{I}^*\overset{\text{def}}{=}\mathbf{I}\cup\mathbf{A}$ with correspondingly enlarged transition matrix function $\prod^*(\cdot)$ in such a way that if in this process "the time spent in the boundary set **A** be deleted", the resulting shrunk process will be the given (or constructed as the case may be) x, **I**, $\prod$. This idea goes

back to Lévy and has been developed by Neveu and Williams. The algebraic part was already indicated by Feller and in the one-atom case by Reuter. We shall show here how the construction of the latter authors can be extended to our case by a simple modification of Theorem 16.1. This is effected by a more general application of the Lemma of § 14, which is indeed a kind of analytical shadow of Lévy's idea of alloting time to fictitious states. The development of this paper makes it clear that this allotment may be done for each boundary atom separately.

Let k denote a general element of $\mathbf{I}^*$ defined above. For each a in $\mathbf{A}$, choose a number $p^a \geqslant 0$ and introduce the new quantities (dropping "$\hat{\ }$" as before):

$$z^a_b \overset{\text{def}}{=} \delta^{ab},$$

$$\eta^a_b(\lambda) \overset{\text{def}}{=} p^a \delta^{ab}.$$

Thus c^a, σ^{ab} and u^{ab} are not affected for $a \neq b$, but if we denote new quantities by affixing "$*$" to the corresponding old ones, we have

$$u^{*aa}(\lambda) = u^{aa}(\lambda) + p^a\lambda,$$

so that

$$U^*(\lambda) = U(\lambda) + p\lambda I,$$

$$\langle \lambda\eta^*(\lambda), 1\rangle = \langle \lambda\eta(\lambda), 1\rangle + p^a\lambda,$$

$$E^{*a}(\lambda) = \frac{1}{\delta^a + p^a\lambda + u^{aa}(\lambda)},$$

$$M^*(\lambda)^{-1} = M(\lambda)^{-1} + p\lambda I,$$

$$\Theta^*(\lambda, \mu) = \Theta(\lambda, \mu) + pI.$$

It follows from (16.19) that

$$\begin{aligned} U^*(\lambda) - U^*(\mu) &= (\lambda - \mu)\,\Theta^*(\lambda, \mu) = U(\lambda) - U(\mu) + p(\lambda - \mu)\, I \\ &= M(\lambda)^{-1} - M(\mu)^{-1} + p(\lambda - \mu)\, I = M^*(\lambda)^{-1} - M^*(\mu)^{-1}, \end{aligned}$$

which is the new (16.19) needed for the verification of the resolvent equation. The stochasticity is verified exactly as before in (16.21) with the appropriate quantities starred.

It should be pointed out that unless one begins with the canonical decomposition, the extended process constructed above will not yield the correct information on the boundary behavior of the original process. For instance, the transition of the new states in $\mathbf{A}$ among themselves will not be controlled by the jump matrix $(F^{ab}(\infty))$, as it should be, but rather by Ω which may not be the same matrix when the decomposition is not canonical. This leads to the subject matter of the next two sections.

§ 17. Etude approfondie

In this section we begin by giving probabilistic meanings to the quantities, E^a, σ^{ab}, $a \neq b$, which were derived analytically in § 14. Using these we shall be able to analyze in more depth the functions $F^{ab}(\cdot)$ and in particular their limit values $F^{ab}(0)$ and $F^{ab}(\infty)$. These results are also requisite for the problems dealt with in the next section.

A fundamental new random variable, NOT an optional one, will now be introduced.

Definition 17.1. For each boundary atom a, let

$$\gamma^a(\omega) = \sup\{S^a(\omega) \cap (0, \beta^a(\omega))\}.$$

This is the "last exit time from a before switch (changing banners)".

Clearly we have, using the definition in (14.23):

$$\mathbf{P}^a\{\gamma^a < +\infty\} = \delta^a.$$

If a is nonsticky, γ^a is just one of the sequence of times, finite in every finite interval, when the process reaches the boundary at a. If a is sticky, it follows as in the Corollary to Theorem 12.4 that $\gamma^a \in \overline{S^a} - S^a$, so that in particular the process is by definition neither at a nor indeed at any other boundary point. Intuitively, $x(\gamma^a - 0)$ is at an "inaccessible boundary" not definable by means of the jump chain alone. Its behavior is best understood by analogy with the last exit time from an instantaneous state (see [1; Addenda] and [2]). A more comprehensive new boundary theory must surely cover this situation but for our present purposes it is opossible to circumvent this difficulty. The basic nature of γ^a is clearly reflected in the theorem below.

THEOREM 17.1. *We have for each a and $a \neq b$:*

$$\mathbf{P}^a\{\gamma^a \leqslant t\} = E^a(t); \tag{17.1}$$

$$\mathbf{P}^a\{\gamma^a < \beta^a;\ \gamma^a \in ds;\ \beta^a \in dt;\ x(\beta^a) = b\} = E^a(ds)\theta^{ab}(t-s)\,dt. \tag{17.2}$$

Remark. The meaning of the differentials will be explained in the proof. Note also that in (17.2), since $\beta^a > 0$, "$x(\beta^a) = b$" is well-defined, but later when β^a may be zero we must write "$\beta^a \in \overline{S^b}$" instead, as in the proof of Theorem 15.1.

Proof. We begin with the relation, valid for each $t \geqslant 0$:

$$\mathbf{P}^a\{\gamma^a > t\} = \sum_i \mathbf{P}^a\{\beta^a > t;\ x(t) = i;\ \alpha_t < \infty;\ x(\alpha_t) = a\}. \tag{17.3}$$

For we have $\gamma^a \leqslant \beta^a$; and on the set $\{\beta^a > t\}$, if $\alpha_t = +\infty$ then $S^a \cap (t, \infty) = \emptyset$; while if $\alpha_t < \infty$

and $x(\alpha_t) \neq a$, then $\alpha_t = \beta^a$ and $S^a \cap (t, \beta^a) = \emptyset$. Hence either case implies $\gamma^a \leqslant t$. This proves (17.3). Using (14.33), the right member of (17.3) may be written as

$$\langle \varrho^a(t), L^a(\infty) \rangle = \varrho^{aa}(t) = 1 - E^a(t), \tag{17.4}$$

proving (17.1). A similar argument establishes the more specific relation

$$\begin{aligned} \mathbf{P}^a\{\gamma^a < s < t < \beta^a < t'; x(\beta^a) = b\} &= \langle \varrho^a(s), L^b(t'-s) - L^b(t-s) \rangle \\ &= \int_0^s \langle \eta^a(s-u), L^b(t'-s) - L^b(t-s) \rangle \, dE^a(u), \end{aligned} \tag{17.5}$$

the second equation by (14.17). Since the last-written integrand above is

$$\int_{t-s}^{t'-s} \langle \eta^a(s-u), l^b(v) \rangle \, dv = \int_{t-s}^{t'-s} \theta^{ab}(s-u+v) \, dv = \sigma^{ab}(t'-u) - \sigma^{ab}(t-u)$$

by (14.20) and (14.21), the last member in (17.5) reduces to

$$\int_{[0,s)} E^a(du) \int_{[t,t')} \theta^{ab}(r-u) \, dr.$$

It follows that for $0 \leqslant s < s' < t < t'$,

$$\mathbf{P}^a\{s \leqslant \gamma^a < s' < t \leqslant \beta^a < t'; x(\beta^a) = b\} = \int_{[s,s')} E^a(du) \int_{[t,t')} \theta^{ab}(r-u) \, dr,$$

which is what is meant by (17.2).

Corollary 1. *We have*

$$\mathbf{P}^a\{\gamma^a < \beta^a \leqslant t; x(\beta^a) = b\} = \int_0^t [\sigma^{ab}(0) - \sigma^{ab}(t-s)] \, dE^a(s); \tag{17.6}$$

$$\mathbf{P}^a\{\gamma^a = \beta^a \leqslant t; \beta^a \in \overline{S^b}\} = E^a(t) \, [F^{ab}(\infty) - \sigma^{ab}(0)]. \tag{17.7}$$

Proof. It follows from (17.2) and Fubini's theorem on product measure that the left member of (17.6) is equal to

$$\int_{[0,t)} E^a(ds) \int_{[s,t)} \theta^{ab}(t-s) \, ds$$

which is the right member of (17.6). Now we have by Definition 14.1 and (14.35) that

$$\mathbf{P}^a\{\beta^a \leqslant t; \beta^a \in \overline{S^b}\} = F^{ab}(t) = \int_0^t [F^{ab}(\infty) - \sigma^{ab}(t-s)] \, dE^a(s). \tag{17.8}$$

Subtracting (17.6) from (17.8) we obtain (17.7). Q.e.d.

It is essential to understand the meaning of the probability in (17.7). How is $\gamma^a = \beta^a$ possible? This happens if and only if

$$\forall \delta > 0: \quad S^a \cap (\gamma^a - \delta, \gamma^a] \neq \emptyset; \quad S^b \cap (\gamma^a, \gamma^a + \delta) \neq \emptyset,$$

(a situation envisaged in the Remark on p. 39 of [I]). If a is nonsticky, then $x(\gamma^a) \in S^a$; if a is sticky, we have already noted that $x(\gamma^a) \in \overline{S^a} - S^a$. In either case the second relation above is possible only if b is sticky, by Theorem 12.5. Needless to say all the assertions above are true with probability one only. We have incidentally discovered an important number, now to be defined.

Definition 17.2. For $a \neq b$, let

$$d^{ab} = F^{ab}(\infty) - \sigma^{ab}(0). \tag{17.9}$$

COROLLARY 2. *If a is not a recurrent trap, then for every $b \neq a$,*

$$d^{ab} = \mathbf{P}^a\{\gamma^a = \beta^a \in \overline{S^b} \mid \gamma^a\}. \tag{17.10}$$

In particular, $d^{ab} = 0$ for every a if b is nonsticky.

Proof. If a is not a recurrent trap, we know from Theorem 14.4 that $E^a(\infty) = 1$. Letting $t \uparrow \infty$ in (17.7), we obtain

$$d^{ab} = \mathbf{P}^a\{\gamma^a = \beta^a < \infty;\ \beta^a \in \overline{S^b}\}.$$

Substituting this back into the right member of (17.7), and comparing with (17.1), we obtain (17.10).

COROLLARY 3. *For every a and b, $a \neq b$, we have*

$$d^{ab} = \lim_{t \downarrow 0} \frac{F^{ab}(t)}{E^a(t)}. \tag{17.11}$$

Proof. Observe first that $E^a(t) > 0$ for $t > 0$, from Corollary 1 to Theorem 14.7. We have by (14.34) and (14.36)

$$F^{ab}(\infty) = \frac{1}{E^a(t)} \left\{ \int_0^t \sigma^{ab}(t-s)\, dE^a(s) + F^{ab}(t) \right\}.$$

Letting $t \downarrow 0$ the corollary follows, since $\sigma^{ab}(t) \nearrow \sigma^{ab}(0)$.

COROLLARY 4. *For every $a \neq b$, and on the set $\{\gamma^a < t\}$, we have*

$$\mathbf{P}^a\{t < \beta^a < \infty;\ x(\beta^a) = b \mid \gamma^a\} = \sigma^{ab}(t - \gamma^a); \tag{17.12}$$

$$\mathbf{P}^a\{\gamma^a<\beta^a<\infty;\ x(\beta^a)=b\,|\,\gamma^a\}=\sigma^{ab}(0); \tag{17.13}$$

$$\mathbf{P}^a\{\beta^a\in\overline{S^b}\,|\,\gamma^a\}=F^{ab}(\infty); \tag{17.14}$$

$$\mathbf{P}^a\{\beta^a=+\infty\,|\,\gamma^a\}=\varrho^a_*(\infty). \tag{17.15}$$

Proof. By (17.2)

$$\mathbf{P}^a\{\gamma^a\leqslant s<t<\beta^a;\ x(\beta^a)=b\}=\int_0^s\sigma^{ab}(t-u)\,dE^a(u).$$

This being true for fixed t and arbitrary $s<t$, we have (17.12). Next, we write for each $t\geqslant 0$:

$$F^{ab}(\infty)=\mathbf{P}^a\{\beta^a\leqslant t;\ \beta^a\in\overline{S^b}\}+\mathbf{P}^a\{\gamma^a\leqslant t<\beta^a;\ x(\beta^a)=b\}+\mathbf{P}^a\{t<\gamma^a;\ x(\beta^a)=b\}. \tag{17.16}$$

The first term on the right side is $F^{ab}(t)$; the second is by (17.2) equal to

$$\int_0^t\sigma^{ab}(t-s)\,E^a(ds);$$

hence the third is equal to

$$F^{ab}(\infty)-F^{ab}(t)-\int_0^t\sigma^{ab}(t-s)\,E^a(ds)=[1-E^a(t)]\,F^{ab}(\infty)$$

by (14.35). This means, by (17.1):

$$\mathbf{P}^a\{\gamma^a>t;\ \beta^a\in\overline{S^b}\}=\mathbf{P}^a\{\gamma^a>t\}\,\mathbf{P}^a\{\beta^a\in\overline{S^b}\}$$

so that there is independence and (17.14) follows. From this (17.15) follows by (14.4). Finally subtracting (17.10) from (17.14) we obtain (17.13) which is a limiting form of (17.12).

The next theorem is the completed version of Theorem 5.3 in [I]. The proof there is analytic and leaves one important point unsettled, concering $L^{ab}(0+)$, which is specifically mentioned on p. 42 of [I]. The difficulty is resolved here since the relevant sample function behavior has now been clarified and the new proof, given in perhaps excessive detail here owing to past failure, is probabilistic. We put $F^{aa}(0)=0$ for each a below.

THEOREM 17.2. *For each a and b in $\mathbf{A}$ and $0<t\leqslant+\infty$ the limits below exist:*

$$L^{ab}(t)\overset{\text{def}}{=}\lim_{s\downarrow 0}\langle\xi^a(s),L^b(t-s)\rangle, \tag{17.17}$$

$$\tilde{L}^{ab}(t)\overset{\text{def}}{=}\lim_{s\downarrow 0}\langle\varrho^a(s),L^b(t-s)\rangle; \tag{17.18}$$

and we have $$L^{ab}(t)-\tilde{L}^{ab}(t)=F^{ab}(0)=L^{ab}(0+)\geqslant\tilde{L}^{ab}(0+)=0. \tag{17.19}$$

Furthermore, if a is sticky, then

$$L^{ab}(t)\equiv\tilde{L}^{ab}(t)\equiv\delta^{ab}. \tag{17.20}$$

Proof. For convenience' sake, let us state the following lemma which is trivial to prove as soon as it is formulated but useful also in other similar circumstances.

LEMMA. *Let $\{\Lambda_s,\ s\geqslant 0\}$ be a family of sets in $\mathfrak{F}$, and Λ also in $\mathfrak{F}$. Suppose that there exists a sequence of sets Ω_n in $\mathfrak{F}$ such that $\Omega_n\nearrow\Omega$ and such that*

$$\Lambda_s\cap\Omega_n=\Lambda\cap\Omega_n$$

whenever $s<\varepsilon_n$ where $\varepsilon_n\searrow 0$, then

$$\lim_{s\downarrow 0}\mathbf{P}(\Lambda_s)=\mathbf{P}(\Lambda).$$

To prove Theorem 17.2 we prove first (17.20). If a is sticky, for each $t>0$ and a.e. ω, there exists an integer n_0 such that:

$$\forall n\geqslant n_0:\ \alpha_{n^{-1}}(\omega)<t,\ x(\alpha_{n^{-1}})=a.$$

Let $n_0(\omega)$ be the least such integer and put $s_0(\omega)=\alpha_{n_0(\omega)^{-1}}$. Then $s_0(\omega)$ is a random variable satisfying for a.e. ω:

(i) $$0<s_0(\omega)<t,$$

(ii) $$x(s_0(\omega))=a,$$

(iii) $$\bigcup_{b\neq a}S^b(\omega)\cap(0,s_0(\omega))=\emptyset.$$

The last as a consequence of Definition 12.2 and Theorem 4.6 of [I].

Let
$$\Omega_n\overset{\text{def}}{=}\{\omega:s_0(\omega)>n^{-1}\},$$
$$\Lambda_s^b\overset{\text{def}}{=}\{s<\beta^a;\ \alpha_s\leqslant t;\ \alpha_s\in\overline{S^b}\},\quad s\geqslant 0,\ b\in\mathbf{A}$$

We have $\Lambda_s^a\cap\Omega_n=\Omega\cap\Omega_n$ whenever $s<n^{-1}$; hence by the Lemma above

$$\lim_{s\downarrow 0}\mathbf{P}^a(\Lambda_s^a)=\mathbf{P}^a(\Omega)=1.$$

Now by definition we have

$$\mathbf{P}^a(\Lambda_s^a)=\langle\varrho^a(s),L^a(t-s)\rangle,$$

and consequently we have proved that $\tilde{L}^{aa}(t)\equiv 1$. Since

$$\sum_{b\in\mathbf{A}}\tilde{L}^{ab}(t)\leqslant\lim_{s\downarrow 0}\langle\varrho^a(s),1\rangle\leqslant 1,$$

this implies $\tilde{L}^{ab}(t)\equiv 0$ for $b\neq a$. Since $\xi^a(\cdot)\geqslant\varrho^a(\cdot)$, the first equation in (17.20) now follows. We have proved (17.17)–(17.20) for a sticky a.

From now on in this proof let a be nonsticky. Let

$$\Delta^c\overset{\text{def}}{=}\{\omega:0\in\overline{S^c(\omega)}\},\quad\Delta'=\Omega\setminus\bigcup_{c\in\mathbf{A}}\Delta^c.$$

On $\Delta', \beta^a \geqslant \alpha > 0$ a.e. and if $s < \alpha(\omega)$, then $\alpha_s(\omega) = \alpha(\omega)$. Hence if $\Omega'_n = \{\omega : \alpha(\omega) > n^{-1}\}$, we have

$$\Omega'_n \cap \Delta' \cap \Lambda^b_s = \Omega'_n \cap \Delta' \cap \Lambda^b_0$$

whenever $s < n^{-1}$. Hence as before

$$\lim_{s \downarrow 0} \mathbf{P}^a(\Delta' \cap \Lambda^b_s) = \mathbf{P}^a(\Delta' \cap \Lambda^b_0).$$

Since a is nonsticky, under $\mathbf{P}^a$ we have $\alpha^a > 0$ as well as $\beta^a > 0$ on Λ^b_s for $s > 0$, so that

$$\mathbf{P}^a(\Delta' \cap \Lambda^b_s) = \mathbf{P}^a(\Lambda^b_s) = \langle \varrho^a(s), L^b(t-s) \rangle.$$

Together with the preceding relation this proves

$$\tilde{L}^{ab}(t) = \mathbf{P}^a(\Delta' \cap \Lambda^b_0).$$

Next, let $$M^b_s \stackrel{\text{def}}{=} \{\alpha_s \leqslant t, x(\alpha_s) = b\}, \quad s > 0.$$

Then we have by definition

$$\langle \xi^a(s), L^b(t-s) \rangle = \mathbf{P}^a(M^b_s) = \sum_{c \in \mathbf{A}} \mathbf{P}^a(\Delta^c \cap M^b_s) + \mathbf{P}^a(\Delta' \cap M^b_s).$$

Since $\Delta' \cap (M^b_s \backslash \Lambda^b_s) \subset \Delta' \cap \{\beta^a \leqslant s\} \searrow \emptyset$ as $s \downarrow 0$, we have

$$\lim_{s \downarrow 0} \mathbf{P}^a(\Delta' \cap M^b_s) = \lim_{s \downarrow 0} \mathbf{P}^a(\Delta' \cap \Lambda^b_s),$$

hence the last term above converges to $\tilde{L}^{ab}(t)$ as just shown. For each sticky c, we have by Lemma 1 of § 15:

$$\lim_{s \downarrow 0} \mathbf{P}^a(\Delta^c \cap M^b_s) = F^{ac}(0) \lim_{s \downarrow 0} \mathbf{P}^c(M^b_s) = F^{ac}(0)\, L^{cb}(t) = F^{ac}(0)\, \delta^{cb},$$

the second equation above being an application of (17.17) and the third of (17.20) both with $a = c$. Combining these, we obtain

$$L^{ab}(t) = \sum_{c \in \mathbf{A}} F^{ac}(0)\, \delta^{cb} + \tilde{L}^{ab}(t) = F^{ab}(0) + \tilde{L}^{ab}(t).$$

This proves the first equation in (17.19); to prove the second, we observe that by definition

$$L^{ab}(t) \leqslant \mathbf{P}^a\{S^b \cap (0, t) \neq \emptyset\}$$

and consequently $L^{ab}(0+) \leqslant F^{ab}(0) \leqslant L^{ab}(0+)$. The rest follows. Q.e.d.

The following result is an essential sharpening of the preceding theorem and is also the major step towards the identification problem in the next section. It seems to depend almost precariously on the finer properties of boundary atoms.

THEOREM 17.3. *For each a, we have*

$$\lim_{s\downarrow 0}\frac{1-\langle \xi^a(s), L^a(\infty)\rangle}{E^a(s)}=\delta^a; \tag{17.21}$$

and if $\delta^a=1$, for $0<t\leqslant+\infty$:

$$\lim_{s\downarrow 0}\frac{1-\langle \xi^a(s), L^a(t)\rangle}{1-\langle \xi^a(s), L^a(\infty)\rangle}=1+\sigma^{aa}(t). \tag{17.22}$$

For each $c\neq a$ and $0<t\leqslant+\infty$:

$$\lim_{s\downarrow 0}\frac{\langle \xi^a(s), L^c(t)\rangle}{1-\langle \xi^a(s), L^a(\infty)\rangle}=F^{ac}(\infty)-\sigma^{ac}(t); \tag{17.23}$$

in particular

$$\lim_{s\downarrow 0}\frac{\langle \xi^a(s), L^c(\infty)\rangle}{1-\langle \xi^a(s), L^a(\infty)\rangle}=F^{ac}(\infty). \tag{17.24}$$

Proof. We have by (15.1):

$$\langle \xi^a(s), L^c(t)\rangle=\langle \varrho^a(s), L^c(t)\rangle+\sum_{b\neq a}\int_0^s\langle \xi^b(s-u), L^c(t)\rangle\, dF^{ab}(u). \tag{17.25}$$

The first term on the right is equal to

$$\int_0^s\langle \eta^a(s-u), L^c(t)\rangle\, dE^a(u)=\int_0^s[\sigma^{ac}(s-u)-\sigma^{ac}(s-u+t)]\, dE^a(a).$$

Recalling (17.11), we have as $s\downarrow 0$:

$$F^{ab}(s)\sim E^a(s)\, d^{ab}.\ ^{(1)}$$

Substituting the last two equations in (17.25), we have as $s\downarrow 0$:

$$\langle \xi^a(s), L^c(t)\rangle\sim E^a(s)\left\{\sigma^{ac}(0)-\sigma^{ac}(t)+\sum_{b\neq a}d^{ab}L^{bc}(t)\right\}.$$

But by Corollary 2 to Theorem 17.1, $d^{ab}=0$ unless b is sticky, and if so $L^{bc}(t)=\delta^{bc}$ by (17.20); hence

$$\sum_{b\neq a}d^{ab}L^{bc}(t)=\sum_{b\neq a}d^{ab}\delta^{bc}=\begin{cases}d^{ac} & \text{if}\quad a\neq c,\\ 0 & \text{if}\quad a=c.\end{cases}$$

Consequently if $a\neq c$, we have by (17.9):

$$\langle \xi^a(s), L^c(t)\rangle\sim E^a(s)\left\{\sigma^{ab}(0)-\sigma^{ac}(t)+d^{ac}\right\}=E^a(s)\left\{F^{ac}(\infty)-\sigma^{ac}(t)\right\}. \tag{17.26}$$

(1) $u(s)\sim v(s)$ means $\lim_{s\downarrow 0}u(s)/v(s)=1$.

Next, putting $a=c$ and $t=+\infty$ in (17.25), we have

$$\langle \xi^a(s), L^a(\infty)\rangle = \int_0^s \sigma^{aa}(s-u)\,dE^a(u) + \sum_{b\neq a}\int_0^s \langle \xi^b(s-u), L^a(\infty)\rangle\, dF^{ab}(u).$$

The first term on the right side is equal to $1-\delta^a E^a(s)$ by Corollary 1 to Theorem 14.7. Hence as $s \downarrow 0$ we have as before:

$$1-\langle \xi^a(s), L^a(\infty)\rangle \sim E^a(s)\,[\delta^a - \sum_{b\neq a} d^{ab} L^{ba}(\infty)] = E^a(s)\,\delta^a.$$

This proves (17.21), and together with (17.26) proves (17.23). If $\delta^a = 1$ a similar argument yields

$$1-\langle \xi^a(s), L^a(t)\rangle \sim E^a(s)\,[1+\sigma^{aa}(t)].$$

Hence (17.22) follows from this and (17.21).

Remark. (17.24) may be written as

$$\lim_{\substack{s\downarrow 0\\ s>0}} \frac{\mathbf{P}^a\{x(\alpha_s)=c\}}{1-\mathbf{P}^a\{x(\alpha_s)=a\}} = F^{ac}(\infty),$$

or even more suggestively as

$$\lim_{\substack{s\downarrow 0\\ s>0}} \frac{\mathbf{P}^a\{x(\alpha_s)=b\}}{\mathbf{P}^a\{x(\alpha_s)=c\}} = \frac{F^{ab}(\infty)}{F^{ac}(\infty)}$$

provided $F^{ac}(\infty)>0$. Is either of these relations "intuitively obvious"?

§ 18. Identification

In this section we consider the following problem. Given $\prod = (p_{ij}(\cdot))$, how to "find" the quantities L^a, ξ^a, F^{ab}, ϱ^a, e^a, E^a, η^a, σ^{ab}, introduced earlier in the paper? These have all been defined in terms of the process $\{x_t\}$, but how can they be expressed, or at least, determined by means of $\prod$? The words "expression" and "determination" are of course themselves subject to interpretation but our results below will be stated in specific ways. This problem, to be called that of "identification", has its proper interest, but we shall also use its solution to answer the questions raised in § 11, concerning the tracking down of the canonical decomposition and the construction of stopped processes.

Let Λ be a set in the Borel field $\mathfrak{F}^0$ generated by $\{x_t, 0\leq t<\infty\}$. Since the basic probability measure $\mathbf{P}$ on $\mathfrak{F}^0$ is uniquely determined by its values on cylinder sets of the form $\{x(t_n)=j_n,\ 0\leq n\leq l\}$, $t_n\geq 0$, $j_n\in \mathbf{I}$, it is determined by $\prod$ and the initial distribution of the

process. In particular, $\mathbf{P}_i\{...\}$ for each $i \in \mathbf{I}$ is determined by $\prod$ alone. However, many interesting sets Λ must in practice be defined in terms of a "nice version" of $\{x_t\}$ (see [1; § II.7]); hence the following simple observation is necessary.

LEMMA. *Let two arbitrary stochastic processes $\{x_t\}$ and $\{\tilde{x}_t\}$ defined on the probability triple $(\Omega, \mathfrak{F}, \mathbf{P})$ be standard modifications of each other, then the Borel fields $\mathfrak{F}^0$ and $\tilde{\mathfrak{F}}^0$ generated by them are identical provided they be both augmented (by all $\mathbf{P}$-null sets).*

Proof. If $\Lambda \in \mathfrak{F}^0$, it is well-known that there is a countable set $\{t_n\}$, $n \in \mathbf{N}$, such that $\Lambda \in \mathfrak{F}\{x(t_n),\ n \in \mathbf{N}\}$. For each n, $\mathbf{P}\{x(t_n) = \tilde{x}(t_n)\} = 1$ and so $\mathbf{P}\{\forall n \in \mathbf{N}: x(t_n) = \tilde{x}(t_n)\} = 1$. It follows that Λ differs from a set in $\mathfrak{F}\{\tilde{x}(t_n),\ n \in \mathbf{N}\}$ by a null set, hence $\Lambda \in \tilde{\mathfrak{F}}^0$ and the proof is finished.

As a consequence of the lemma, quantities such as those mentioned above are determined by $\prod$ in the sense that for two processes having the same transition matrix $\prod$, these quantities have the same values. However, even in the simplest cases involving no boundary, an explicit expression may be hard to come by, for example the taboo probability ${}_0p_{0j}(t)$ used in the proof of Theorem 14.3. Here indeed lies the great advantage of the probabilistic method, by relying on the sample functions of the process rather than its transition matrix. However, we are now pushing matters in the opposite direction.

To recapitulate from the beginning of the story (see [I; §§ 2–4]), Q is the initial derivative matrix $(p'_{ij}(0))$, supposed to be conservative; given Q, $\Phi(\cdot)$ can be constructed by a purely analytic iteration procedure (given 25 years ago by Feller) as the minimal solution of the Kolmogorov differential equations, both backward and forward. Q also determines (it is trivially equivalent to) the jump matrix $\mathsf{P} = (r_{ij})$ where $r_{ij} = (1 - \delta_{ij}) q_{ij} q_i^{-1}$. Now it is necessary to confront the boundary. Under Assumption B, that the passable part of the boundary be atomic, each boundary atom a, $a \in \mathbf{A}$, corresponds to an essentially uniquely determined atomic almost closed set A^a of $\mathbf{I}$, and we have (cf. [1; § I.17]):

$$L_i^a(\infty) = \lim_{n\to\infty} \mathsf{P}_i\{\chi_n \to a\} = \lim_{n\to\infty} \sum_{j \in A^a} r_{ij}^{(n)},$$

where $\{\chi_n\}$ is the jump chain whose probability behavior is controlled by its transition matrix P. The above is substantially Feller's definition of a sojourn solution [6]. There is another way of identifying $\{L^a(\infty),\ a \in \mathbf{A}\}$: they are the extreme points of the cone of solutions of the equation:

$$Qe = 0$$

with $e \leqslant 1$; see the discussion around Theorem 4.2 of [I]. This is also due to Feller.

Next, ξ^a may be identified by Theorem 4.5 of [I]:

$$\xi_j^a(t) = \lim_{\chi_n \to a} p_{\chi_n j}(t) \tag{18.1}$$

with probability one, for each j and $t>0$. The preceding formula, like the one above for $L^a(\infty)$, uses the jump chain and its boundary behavior. This seems inevitable as somehow the *fact of approaching* the boundary at a specified atom must be expressed, nevertheless there is the question of the "neatest" way of doing so. One may remark in passing that the Martin boundary theory (see [8]) yields a similar but less precise expression for L^∞ as follows:

$$\frac{L_i^a(\infty)}{L_0^a(\infty)} = \lim_{\chi_n \to a} \frac{\sum_{n=0}^{\infty} r_{i\chi_n}^{(n)}}{\sum_{n=0}^{\infty} r_{0\chi_n}^{(n)}}$$

where 0 is some fixed state.

Given ξ^a and $\prod$ the probability measure $P^a\{...\}$ on the Borel field $\mathfrak{F}\{x_t, t>0\}$ is uniquely determined, since if $0<t_1<...<t_l$, and $j_n \in \mathbf{I}$ we have

$$\mathbf{P}^a\{x(t_n) = j_n, 1 \leqslant n \leqslant l\} = \xi_{j_1}^a(t_1) \prod_{\nu=1}^{l-1} p_{j_\nu j_{\nu+1}}(t_{\nu+1} - t_\nu).$$

This may be extended to include the sets $\{x(0) = j\}$, since

$$\mathbf{P}^a\{x(0) = j\} = \lim_{t \downarrow 0} \mathbf{P}^a\{x(t) = j\}.$$

Analytically, the identification given for $\{\xi^a, a \in \mathbf{A}\}$ is sufficient to yield the following useful uniqueness theorem (cf. Reuter [14]).

THEOREM 18.1. *Suppose that there is a decomposition of the form*

$$\prod(t) = \Phi(t) + \sum_{a \in \mathbf{A}} \int_0^t y^a(t-s)\, dL^a(s), \tag{18.2}$$

where $\prod$, Φ, L *have previous meanings; and either for each* a, y^a *is measurable and* $\geqslant 0$, *or for each* a, $\int_0^\infty e^{-\lambda t} |y^a(t)|\, dt < \infty$ *for* $\lambda>0$. *Then we have for each* a *and almost every* t:

$$y^a(t) = \xi^a(t).$$

If in addition $y^a(\cdot)$ *is right or left continuous, then there is equality above for every* t.

Proof. Taking Laplace transforms, we have

$$\hat{p}_{ij}(\lambda) = \hat{f}_{ij}(\lambda) + \sum_{b \in \mathbf{A}} \hat{L}_i^b(\lambda)\, \hat{y}_j^b(\lambda). \tag{18.3}$$

The alternative condition on the y^a's ensures that the Laplace transforms of the convolution in (18.2) is the product indicated under the sum in (18.3), if we observe that $0 \leqslant \prod(t) - \Phi(t) \leqslant 1$. Now substitute χ_n for i in (18.3) and let $\chi_n \to a$ as in (18.1). As a consequence of Theorem 4.1 of [I], we have

$$\lim_{\chi_n \to a} \hat{L}^b_{\chi_n}(\lambda) = \delta^{ab}, \qquad \lim_{\chi_n \to a} \hat{f}_{\chi_n j}(\lambda) = 0;$$

this together with the Laplace transform of (18.1) yields at once

$$\hat{\xi}^a_j(\lambda) = 0 + \sum_b \delta^{ab} \hat{y}^b_j(\lambda) = \hat{y}^b_j(\lambda).$$

The assertions of the theorem follow from the uniqueness theorem for Laplace transforms and the continuity of $\xi^a_j(\cdot)$. Q.e.d.

Given ξ^a and $L^a(\infty)$, $a \in \mathbf{A}$, the other quantities can be expressed purely analytically without further intervention of the boundary. This will be exhibited, repeating previous formulae to avoid excessive cross references, in the following summing-up.

THEOREM 18.2. (1) *a is not a recurrent trap: $\delta^a = 1$.*

$$L^a(t) = [I - \Phi(t)]\, L^a(\infty);$$

$$1 + \sigma^{aa}(t) = \lim_{s \downarrow 0} \frac{1 - \langle \xi^a(s), L^a(t) \rangle}{1 - \langle \xi^a(s), L^a(\infty) \rangle};$$

$E^a(\cdot)$ is the unique solution of the integral equation

$$\forall t > 0{:}\ \int_0^t [\delta^a + \sigma^{aa}(t-s)]\, dE^a(s) = 1;$$

$$F^{ab}(\infty) = \lim_{s \downarrow 0} \frac{\langle \xi^a(s), L^b(\infty) \rangle}{1 - \langle \xi^a(s), L^a(\infty) \rangle}, \quad a \neq b;$$

$$\sigma^{ab}(t) = F^{ab}(\infty) - \lim_{s \downarrow 0} \frac{\langle \xi^a(s), L^b(t) \rangle}{1 - \langle \xi^a(s), L^a(\infty) \rangle}; \quad a \neq b;$$

$$\begin{aligned} \varrho^a_j(t) &= \xi^a_j(t) - \sum_{b \neq a} \int_0^\infty \xi^b_j(t-s)\, dF^{ab}(s) \quad && \text{if} \quad j \in \mathbf{I}^a, \\ &= 0 && \text{if} \quad j \notin \mathbf{I}^a; \end{aligned}$$

$$e^a = \int_0^\infty \varrho^a(t)\, dt;$$

$$\eta^a(t) = \frac{d}{dt} e^a [I - \Phi(t)];$$

(2) *a is a recurrent trap:* $\delta^a = 0$.

$L^a(t)$ *as above;*

$e_j^a = \int_0^\infty {}_0p_{0j}(t)\,dt$, 0 *fixed in* $\mathbf{I}^a$, $j \in \mathbf{I}^a$;

$\eta^a(t)$ *as above with the e just defined;*

$\sigma^{aa}(t) = \langle \eta^a(t), L^a(\infty) \rangle$;

$E^a(\cdot)$ *as above;*

$\varrho^a \equiv \xi^a$; $F^{ab}(\cdot) \equiv \sigma^{ab}(\cdot) \equiv 0$; $a \neq b$.

As illustrations, the preceding theorem enables us to solve the following problems. The point is we can go from one set of analytical data to another without introducing new quantities.

Problem 1. Given a process constructed by Theorem 16.1, with $z^a = L^a(\infty)$, $a \in \mathbf{A}$; to find its canonical decomposition as given in Theorem 15.2.

The solution is immediate since the construction produces ξ^a, $a \in \mathbf{A}$, as well as $\prod$, from which we determine δ^a by putting $\delta^a = 0$ if and only if

$$\langle \xi^a(s), 1 - L^a(\infty) \rangle \equiv 0,$$

otherwise $\delta^a = 1$. Now in either case of Theorem 18.2, we obtain ϱ^a, E^a, η^a, F^{ab}, thus retrieving the canonical decomposition (15.6).

Problem 2. Given $L^a(\infty)$ and $\xi^a(\cdot)$, $a \in \mathbf{A}$, constructed or otherwise; to find the transition matrix of the process stopped at $\mathbf{A}_1$, a subset of $\mathbf{A}$.

If $\{x(t)\}$ is the original full process, the stopped process $\{\tilde{x}(t)\}$ is defined as follows. Recalling the definition of α^a in § 13, we put $\alpha^{\mathbf{A}_1} = \inf_{a \in \mathbf{A}_1} \alpha^a$, and

$$\tilde{x}(t) = \begin{cases} x(t) & \text{if } \quad 0 \leqslant t \leqslant \alpha^{\mathbf{A}_1}, \\ \theta & \text{if } \qquad t > \alpha^{\mathbf{A}_1}; \end{cases}$$

where θ is the adjoined absorbing state. For $\mathbf{A}_1 = \mathbf{A} - \{a\}$, the stopped process is just the a-process defined near the beginning of § 14, with the transition matrix $\prod^a$ given there. The general solution to Problem 2 should now be obvious. First we find ϱ^a and F^{ab} by Theorem 18.2; then we set

$$\tilde{\xi}^a = \varrho^a + \sum_{b \in \mathbf{A}_1} \int_0^t \tilde{\xi}^b(t-s)\, dF^{ab}(s), \quad a \in \mathbf{A}_1,$$

or its Laplace transform

$$\check{\xi}^a(\lambda) = \varrho^a(\lambda) + \sum_{b \in \mathbf{A}_1} \hat{F}^{ab}(\lambda)\check{\xi}^b(\lambda), \quad a \in \mathbf{A}_1.$$

Let the restriction of $\hat{F}(\lambda)$ to $\mathbf{A}_1 \times \mathbf{A}_1$ be $\hat{F}_1(\lambda)$, then $I - \hat{F}_1(\lambda)$ is invertible as in the proof of Theorem 15.2. Hence we may solve for $\check{\xi}$:

$$\check{\xi}(\lambda) = [I - \hat{F}_1(\lambda)]^{-1}\hat{\varrho}(\lambda);$$

and

$$\check{\prod}(\lambda) = \Phi(\lambda) + \sum_{a \in \mathbf{A}_1} \hat{L}^a(\lambda)\check{\xi}^a(\lambda)$$

is the Laplace transform of the required stopped transition matrix (without completion).

The procedure given in Theorem 18.2 is somewhat tedious to follow in practice (see the Example at the end of this section). Quicker results can be obtained by uncovering a certain linear transformation which reduces a given decomposition to the canonical form. In what follows we shall again omit the "$\hat{\ }$" on Laplace transforms.

THEOREM 18.3. *Let* $0 < \lambda < \infty$; *and*

$$\xi(\lambda) = M(\lambda)\eta(\lambda),$$

where

$$M(\lambda) = [I - \Omega + U(\lambda)]^{-1}$$

is any construction given by the first part of Theorem 16.1; and let $\tilde{M}$, $\tilde{\eta}$, $\tilde{\Omega}$, $\tilde{U}$ *be any other one. Then there exists a constant invertible matrix* C *such that*

$$\tilde{\eta}(\cdot) = C\eta(\cdot). \tag{18.4}$$

Proof. We have

$$M(\lambda)\eta(\lambda) = \xi(\lambda) = \tilde{M}(\lambda)\tilde{\eta}(\lambda). \tag{18.5}$$

Recalling the resolvent equation, valid for $0 < \lambda < \infty$, $0 < \mu < \infty$:

$$(\lambda - \mu)\eta(\lambda)\Phi(\mu) = \eta(\mu) - \eta(\lambda)$$

for η, and a similar one for $\tilde{\eta}$, we obtain from (18.5)

$$M(\lambda)[\eta(\mu) - \eta(\lambda)] = \tilde{M}(\lambda)[\tilde{\eta}(\mu) - \tilde{\eta}(\lambda)].$$

Cancelling against (18.5), we have

$$M(\lambda)\eta(\mu) = \tilde{M}(\lambda)\tilde{\eta}(\mu). \tag{18.6}$$

This being true for every $\lambda > 0$, $\mu > 0$, we may fix any $\lambda = \lambda_0 > 0$ above and put

$$C = \tilde{M}(\lambda_0)^{-1}M(\lambda_0)$$

to conclude (18.4).

In Theorems 18.4 and 18.5 below we shall assume that the set of $|\mathbf{A}|$ vectors $\{\eta^a(\mu), a \in \mathbf{A}\}$, each regarded as an element of the vector space $\mathfrak{M}(\mathbf{I})$ mentioned in Definition 13.1 above, is linearly independent for some value of $\mu: 0 < \mu < \infty$. Then it follows at once from the resolvent equation for each $\eta^a(\cdot)$ that it is linearly independent for every such value of μ. Under this assumption the equation (18.6) implies that we have for every λ, $0 < \lambda < \infty$:

$$\tilde{M}(\lambda)^{-1}M(\lambda) = C,$$

where C is as in (18.4).

THEOREM 18.4. *If $I - \Omega$ is invertible for some Ω in Theorem 18.3, then it is invertible for every $\tilde{\Omega}$ and we have*

$$C = (I - \tilde{\Omega})(I - \Omega)^{-1}. \tag{18.7}$$

This situation obtains if and only if all boundary atoms are nonrecurrent.

Proof. By (16.6)–(16.9) we have $\langle \eta^a(\mu), 1\rangle < \infty$ for each $\mu > 0$. Using the resolvent equation for η^a we have for each $\lambda > 0$:

$$u^{ab}(\lambda) = \lambda\langle \eta^a(\mu), z^b\rangle + (\mu - \lambda)\langle \eta^a(\mu), \lambda\Phi(\lambda)z^b\rangle;$$

hence $\lim_{\lambda\downarrow 0} u^{ab}(\lambda) = 0$ for each a and b by (16.3), namely $\lim_{\lambda\downarrow 0} U(\lambda) = 0$. Since $M(\lambda)^{-1} = I - \Omega + U(\lambda)$ we obtain

$$\lim_{\lambda\downarrow 0} M(\lambda)^{-1} = I - \Omega. \tag{18.8}$$

It follows from a well-known proposition of finite matrix theory that (18.8) implies

$$\lim_{\lambda\downarrow 0} M(\lambda) = (I - \Omega)^{-1}$$

in the sense that the existence of one member of the equation implies that of the other and also the equality of both. Hence if $I - \Omega$ is invertible, then

$$\lim_{\lambda\downarrow 0} \tilde{M}(\lambda) = \lim_{\lambda\downarrow 0} M(\lambda)\, C^{-1} = (I - \Omega)^{-1}C^{-1}$$

which implies that $I - \tilde{\Omega}$ is invertible and

$$(I - \tilde{\Omega})^{-1} = (I - \Omega)^{-1}C^{-1}$$

proving (18.7). On the other hand, for the canonical $\Omega = (F^{ab}(\infty))$, $I - \Omega$ is invertible unless there exists $\mathbf{A}_0 \subset \mathbf{A}$ such that $\Omega|_{\mathbf{A}_0}$ is stochastic. This is the case if and only if the boundary atoms in $\mathbf{A}_0$ are recurrent. Thus $I - \Omega$ is invertible if all the boundary atoms are nonrecurrent.

THEOREM 18.5. *If all boundary atoms are nonsticky, we have*

$$C=(I-\tilde{\Omega}+\tilde{U}(\infty))(I-\Omega+U(\infty))^{-1}.$$

This is proved in a similar way as the preceding theorem. We now give an example to show the possibility of distinct decompositions mentioned in § 11.

Example. Let $\mathbf{A}=\{a, b\}$; given z^a and z^b with $z^a+z^b=z$, suppose that η^a, η^b are linearly independent and such that

$$\begin{array}{lll} c^a=0, & u^{aa}(\infty)=+\infty, & u^{ab}(\infty)=\frac{1}{2}, \\ c^b=1, & u^{ba}(\infty)=0, & u^{bb}(\infty)<\infty; \\ \Omega^{ab}=1, & \Omega^{ba}=0. & \end{array}$$

These choices are consistent with the conditions of construction; in particular the numerical values of $u^{ab}(\infty)$ and c^b may be fixed by a proportional constant factor.

We have

$$I-\Omega+U(\lambda)=\begin{pmatrix} 1+u^{aa}(\lambda) & -1+u^{ab}(\lambda) \\ 0 & 1+u^{bb}(\lambda) \end{pmatrix}.$$

$$M(\lambda)=[I-\Omega+U(\lambda)]^{-1}=\begin{pmatrix} E^a(\lambda) & E^a(\lambda)\,E^b(\lambda)\,[1-u^{ab}(\lambda)] \\ 0 & E^b(\lambda) \end{pmatrix},$$

where $E^c(\lambda)=[1+u^{cc}(\lambda)]^{-1}$ for $c=a, b$. Thus

$$M(\lambda)\,U(\lambda)=\begin{pmatrix} E^a(\lambda)\,u^{aa}(\lambda) & E^a(\lambda)\,u^{ab}(\lambda)+E^a(\lambda)\,E^b(\lambda)\,[1-u^{ab}(\lambda)]\,u^{bb}(\lambda) \\ 0 & E^b(\lambda)\,u^{bb}(\lambda) \end{pmatrix}.$$

Noting that

$$\langle\xi^a(\lambda), z^b\rangle=\sum_c M^{ac}(\lambda)\,\langle\eta^c(\lambda), z^b\rangle=\sum_c M^{ac}(\lambda)\,\lambda^{-1}u^{cb}(\lambda)=\lambda^{-1}M(\lambda)\,u(\lambda)\,|^{ab},$$

we can compute from the preceding matrix to obtain

$$\lim_{\lambda\uparrow\infty}\frac{\langle\xi^a(\lambda), z^b\rangle}{1-\langle\xi^a(\lambda), z^a\rangle}=\frac{u^{ab}(\infty)+u^{bb}(\infty)}{1+u^{bb}(\infty)}=u<1,$$

$$\lim_{\lambda\uparrow\infty}=\frac{\langle\xi^b(\lambda), z^a\rangle}{1-\langle\xi^b(\lambda), z^b\rangle}=0.$$

Thus the canonical $\tilde{\Omega}=(F^{ab}(\infty))$ is $\begin{pmatrix} 0 & u \\ 0 & 0 \end{pmatrix}$ not the constructed $\Omega=\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.

Theorem 18.4 is applicable and we obtain

$$C=\begin{pmatrix}1 & -u\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & -1\\ 0 & 1\end{pmatrix}^{-1}=\begin{pmatrix}1 & 1-u\\ 0 & 1\end{pmatrix},$$

$$\tilde{\eta}^a=\eta^a+(1-u)\,\eta^b,\qquad \tilde{\eta}^b=\eta^b.$$

The canonical decomposition is therefore

$$[C(I-\Omega+U(\lambda))]^{-1}\,C\eta(\lambda)=[I-\tilde{\Omega}+\tilde{U}(\lambda)]^{-1}\tilde{\eta}(\lambda).$$

These computations can also be made directly by Theorem 18.2.

§ 19. Complements

The first complement concerns nonsticky atoms, for which many of the results and their derivations can be simplfied. While our general discussion includes both sticky and nonsticky atoms it is worthwhile to examine for a little the simpler case where the intuitive meaning is more easily recognized and the analysis follows a smoother pattern. Indeed the sticky case may be regarded as a suitable limiting phenomenon, closely related to the passage of compound Poisson laws to an infinitely divisible law, which deserves further investigation (cf. in this connection a conjecture of Reuter [14], since proved by Kingman [III]).

There are several ways of direct handling of a nonsticky atom a. One is to begin with the Lemma in § 14 and observe the special form of $E^a(\cdot)$ as indicated in Theorem 9.1 of [I]. A more instructive approach, however, is to begin with the basic interpretation of $E^a(\cdot)$ given in (17.1). Recall from Theorem 12.5 that if a is nonsticky and the process starts at a, then the successive times at which it is reached is a sequence of optional random variables

$$0=\tau_0<\tau_1<\tau_2<\ldots,$$

the number of which being finite or infinite with probability one according as a is non-recurrent or recurrent, by Theorem 12.1. As usual we define all the non-existing τ's as $+\infty$.

It is clear from the meaning of L^{aa} in Theorem 17.2 that $L^{aa}(0+)=0$ for a nonsticky a, and that $L^{aa}(\cdot)$ is a probability distribution function if and only if a is a recurrent trap. If we write $L^{(n)aa}$ for the n-fold convolution of L^{aa} with itself, with $L^{(0)aa}=\varepsilon$ the unit mass at 0, we have by the Strong Markov property:

$$\mathbf{P}^a\{\tau_n<\beta^a<\tau_{n+1};\ \tau_n\leqslant t\}=L^{(n)aa}(t)\,[1-L^{aa}(\infty)].$$

Now it follows from (17.1) that

$$E^a(t)=\sum_{n=0}^{\infty}\mathbf{P}^a\{\tau_n<\beta^a<\tau_{n+1},\tau_n\leqslant t\}=\sum_{n=0}^{\infty}L^{(n)aa}(t)\,[1-L^{aa}(\infty)]. \tag{19.1}$$

It is clear that $E^a(t)$ is just the expected number of times of reaching a in time t. [When a is sticky, $E^a(t)$ can be shown to be the expected "local time at a", but this interpretation seems less convenient than that given in (17.1).] Letting $t \downarrow 0$ in Corollary 1 to Theorem 14.7, we obtain

$$E^a(0)[\delta^a + \sigma^{aa}(0)] = 1$$

so that
$$E^a(0) = 1 - L^{aa}(\infty) = \frac{1}{\delta^a + \sigma^{aa}(0)}.$$

It is easy to express F^{ab} in terms of L^{ab}:

$$F^{ab}(t) = L^{ab}(t) + \int_0^t F^{ab}(t-s)\, dL^{aa}(s), \quad a \neq b,$$

or in terms of Laplace transforms:

$$\hat{F}^{ab}(\lambda) = \frac{\hat{L}^{ab}(\lambda)}{1 - \hat{L}^{aa}(\lambda)}. \tag{19.2}$$

In particular
$$F^{ab}(\infty) = \frac{L^{ab}(\infty)}{1 - L^{aa}(\infty)} = \frac{L^{ab}(\infty)}{E^a(0)}.$$

On the other hand, we have by (17.19), (14.35) and (17.9):

$$F^{ab}(0) = L^{ab}(0+) = E^a(0)\,[F^{ab}(\infty) - \sigma^{ab}(0)] = E^a(0)\, d^{ab}.$$

Next we have, by (17.18), (14.17), and (14.27),

$$L^{ab}(t) = \lim_{s \downarrow 0} \int_0^s \langle \eta^a(s-u), L^b(t) \rangle\, dE^a(u) = E^a(0) \int_0^t \theta^{ab}(s)\, ds.$$

Recalling (5.7) on p. 40 of [I], in our present notation:

$$\zeta_i^a(t) \overset{\text{def}}{=} \mathbf{P}^a\{\alpha > 0; x(t) = i\},$$

it is clear that
$$\varrho_i^a(t) = \zeta_i^a(t) + \int_0^t \varrho_i^a(t-s)\, dL^{aa}(s).$$

Namely, $\zeta_i^a(t)$ represents the probability, starting at a, of $x(t) = i$ without having reached any boundary atom before time t, while $\varrho_i^a(t)$ that of the same without reaching any boundary atom *except* a. In Laplace transforms, the equation above is

$$\hat{\varrho}^a(\lambda) = \frac{\hat{\zeta}^a(\lambda)}{1 - \hat{L}^{aa}(\lambda)}. \tag{19.3}$$

Substituting (19.2) and (19.3) into (15.2), we obtain

$$\hat{\xi}^a(\lambda) = \frac{1}{1-\hat{L}^{aa}(\lambda)}\{\hat{\zeta}^a(\lambda) + \sum_{b\in\mathbf{A}-\{a\}} \hat{L}^{ab}(\lambda)\,\hat{\xi}^b(\lambda)\}; \tag{19.4}$$

or clearing of fractions:

$$\hat{\xi}^a(\lambda) = \hat{\zeta}^a(\lambda) + \sum_{b\in\mathbf{A}} \hat{L}^{ab}(\lambda)\,\hat{\xi}^b(\lambda),$$

which is (5.20) of [I].

It now follows from

$$\frac{\hat{\zeta}^a(\lambda)}{1-\hat{L}^{aa}(\lambda)} = \hat{\varrho}^a(\lambda) = \hat{E}^a(\lambda)\,\hat{\eta}^a(\lambda) = \frac{1-L^{aa}(\infty)}{1-\hat{L}^{aa}(\lambda)}\,\hat{\eta}^a(\lambda) = \frac{E^a(0)}{1-\hat{L}^{aa}(\lambda)}\,\hat{\eta}^a(\lambda)$$

that

$$\zeta^a(\cdot) = E^a(0)\,\eta^a(\cdot).$$

Thus ζ^a is the part of $\varrho^a = E^a * \eta^a$ which arises from the mass of E^a at 0. The possibility of using ζ, which has an easy meaning, rather than η, accounts largely for the simplicity of the nonsticky case.

The general reduction sketched above shows the sense in which the development in § 14 is an essential extension of the "first approach" in [I, § 5] to the case where some atoms may be sticky.

The second complement concerns the "last exit time from a before time t," as distinguished from the "last exit time from a before switch" introduced in Definition 17.1.

Definition 19.1. For each a and $t \geq 0$:

$$\gamma_t^a(\omega) = \sup\{S^a(\omega) \cap [0,t]\} = \sup\{s : 0 \leq s \leq t;\, x(s,\omega) = a\}.$$

This is the obvious extension of the last exit time from an ordinary state i before time t ([1; p. 261] and [2]); see also Corollary to Theorem 12.4.

We have then

$$\mathbf{P}^a\{\gamma_t^a \leq s \leq t < \beta^a;\, x(t) = j\} = \sum_i \varrho_i^a(s)\, f_{ij}(t-s)$$

$$= \int_0^s \sum_i \eta_i^a(s-u)\, f_{ij}(t-s)\, dE^a(u) = \int_0^s \eta_j^a(t-u)\, dE^a(u). \tag{19.5}$$

Summing over j: $\qquad \mathbf{P}^a\{\gamma_t^a \leq s \leq t < \beta^a\} = \int_0^s \eta_*^a(t-u)\, dE^a(u).$

Thus we have in density form, for $0 \leq u \leq t$:

$$\mathbf{P}^a\{\gamma_t^a \in du;\, t < \beta^a\} = \eta_*^a(t-u)\, dE^a(u),$$

and $$\mathbf{P}^a\{\gamma_t^a \in du; t < \beta^a; x(t) = j\} = \frac{\eta_j^a(t-u)}{\eta_*^a(t-u)} \mathbf{P}^a\{\gamma_t^a \in du; t < \beta^a\}.$$

Putting $s = t$ in (19.5), we obtain

$$\varrho_j^a(t) = \int_0^t \eta_j^a(t-u)\, dE^a(u);$$

this then is the meaning of the fundamental formula (14.17).

It is interesting to compare this with Theorem 17.1 by calculating the following more specific probabilities:

$$\mathbf{P}^a\{\gamma_t^a \leq s \leq t < \beta^a; x(t) = j; \alpha_t \leq u; x(\alpha_t) = b\}$$
$$= \sum_i \varrho_i^a(s) f_{ij}(t-s) L_i^b(u-t) = \int_0^s \eta_j^a(t-r) L_j^b(u-t)\, dE^a(r);$$

$$\mathbf{P}^a\{\gamma_t^a \leq s \leq t < \beta^a; x(t) = j; \alpha_t = +\infty\}$$
$$= \sum_i \varrho_i^a(s) f_{ij}(t-s)[1 - L_j(\infty)] = \int_0^s \eta_j^a(t-r)[1 - L_j(u-t)]\, dE^a(r).$$

Summing over j, we have if $t \leq u$:

$$\mathbf{P}^a\{\gamma_t^a \leq s \leq t < \beta^a; \alpha_t \leq u; x(\alpha_t) = b\}$$
$$= \int_0^s [\sigma^{ab}(t-r) - \sigma^{ab}(u-r)]\, dE^a(r) = \int_0^s \int_{t-r}^{u-r} \theta^{ab}(v)\, dv\, dE^a(r),$$

or in density form, valid for each a and b (not necessarily distinct):

$$\mathbf{P}^a\{t < \beta^a; \gamma_t^a \in ds; \alpha_t \in du; x(\alpha_t) = b\} = E^a(ds)\, \theta^{ab}(u)\, du.$$

We have also

$$\mathbf{P}^a\{\gamma_t^a \leq s \leq t < \beta^a; \alpha_t = +\infty\} = \int_0^s \eta_*^a(\infty)\, dE^a(r) = c^a E^a(s).$$

These results explain quantitatively the second sentence of [I].

Finally, we shall give a description of the boundary behavior of sample functions in the general case discussed in this paper, namely under A, B′, C_1 and D. This will be seen to be the completed version of the description given on pp. 45–46 of [I] for the case where all boundary atoms are nonsticky. A comparison of the two will show again how the present approach does complete the previous one.

Beginning at time $\beta_0 = 0$ with the banner (boundary atom) z_0, let the successive times for changing banners be $\beta_0 \leq \beta_1 \leq \ldots \leq \beta_n \leq \ldots$ where the sequence is terminated at the first β_n which is $+\infty$, or continued indefinitely if they are all finite. Note that each β_n may equal

β_{n+1} with positive probability. Let the successive banners be $z_0, z_1, \ldots, z_n, \ldots$ so that any two consecutive ones are distinct and that between time β_n and β_{n+1} it is the banner z_n which is flying. The time-atom process

$$\begin{pmatrix}\beta_0, \beta_1, \ldots, \beta_n, \ldots \\ z_0, z_1, \ldots, z_n, \ldots\end{pmatrix}$$

with state space $\mathbf{A}\times(0, \infty)$ is a Markov process characterized as follows:

$$\mathbf{P}\left\{\begin{matrix} z_{n+1}=b \\ \beta_{n+1}\leqslant t\end{matrix}\middle|\begin{matrix} z_0, \ldots, z_{n-1}, z_n=a \\ \beta_0, \ldots, \beta_{n-1}, \beta_n\end{matrix}\right\}=F^{ab}(t);$$

$$\mathbf{P}\left\{\beta_{n+1}=+\infty\middle|\begin{matrix} z_1, \ldots, z_{n-1}, z_n=a \\ \beta_1, \ldots, \beta_{n-1}, \beta_n\end{matrix}\right\}=\varrho_*^a(\infty).$$

The process $\{z_n, n\geqslant 0\}$ is a discrete parameter Markov chain with $(F^{ab}(\infty))$, $(a, b)\in\mathbf{A}\times\mathbf{A}$, where $F^{aa}(\infty)\equiv 0$ for each a, as transition matrix. The banner process, defined on $[0, \infty)$ to be at z_n in the time interval $[\beta^n, \beta^{n+1})$, is a semi-Markovian process (see Pyke [V]). Between each change of banners, namely in each time interval (β_n, β_{n+1}), the ordinary states i line up under the banner z_n with the following probabilities:

$$\mathbf{P}\left\{x(t)=i, t<\beta_{n+1}\middle|\begin{matrix} z_0, \ldots, z_{n-1}, z_n=a \\ \beta_0, \ldots, \beta_{n-1}, \beta_n=s\end{matrix}\right\}=\varrho_i^a(t-s), \quad 0\leqslant s\leqslant t.$$

Summing over i, we have

$$\mathbf{P}\left\{t<\beta_{n+1}\middle|\begin{matrix} z_0, \ldots, z_{n-1}, z_n=a \\ \beta_0, \ldots, \beta_{n-1}, \beta_n=s\end{matrix}\right\}=\varrho_*^a(t-s), \quad 0\leqslant s\leqslant t.$$

We have therefore

$$\begin{aligned}
&\mathbf{P}\{\beta_n\leqslant t<\beta_{n+1}; z_k=a_k, 1\leqslant k\leqslant n\mid z_0=a_0\} \\
&=(F^{a_0a_1}*F^{a_1a_2}*\ldots*F^{a_{n-1}a_n})(t)-\sum_{a_{n+1}\neq a_n}(F^{a_0a_1}*\ldots*F^{a_{n-1}a_n}*F^{a_na_{n+1}})(t) \\
&=\int_0^t \varrho_*^{a_n}(t-s)\,d(F^{a_0a_1}*\ldots*F^{a_{n-1}a_n})(s);
\end{aligned}$$

and consequently

$$\mathbf{P}^{a_0}\{x(t)=i\}=\sum_{n=0}^{\infty}\sum_{a_1,\ldots,a_n}\int_0^t \varrho_i^{a_n}(t-s)\,d(F^{a_0a_1}*\ldots*F^{a_{n-1}a_n})(s),$$

where for $n=0$, the last written convolution is $\varepsilon(\cdot)$ as usual. This is the precise meaning of the canonical decomposition in the first equation of (15.6). The meaning of the second equation has just been given above.

References

[I]. Chung, K. L., On the boundary theory for Markov chains. *Acta Math.*, 110 (1963), 19–77.

[II]. —— Sur une équation de convolution. *C.R. Acad. Sci. Paris*, 260 (1965), 4665–4667; 6794–6796.

[III]. Kingman, J. F. C., The stochastic theory of regenerative events. *Z. Wahrscheinlichkeitstheorie und verw. Gebiete*, 2 (1964), 180–224.

[IV]. Lévy, P., *Théorie de l'addition des variables aléatoires*. Gauthier-Villars, Paris 1937.

[V]. Pyke, R., Markov renewal processes. *Ann. Math. Statist.*, 32 (1961), 1231–1259.

[VI]. Taussky, O., A recurring theorem on determinants. *Amer. Math. Monthly*, 56 (1949), 672–676.

[VII]. Williams, David, The process extended to the boundary. *Z. Wahrscheinlichkeitstheorie und verw. Gebiete*, 2 (1962), 332–339.

[VIII]. —— On the construction problem for Markov chains. *Z. Wahrscheinlichkeitstheorie und verw. Gebiete*, 3 (1964), 227–246.

Received May 19, 1965

Acta Math. 123 (1969) 225–251.
With kind permission of Springer Science and Business Media.

TO REVERSE A MARKOV PROCESS

BY

K. L. CHUNG and JOHN B. WALSH (¹)

Stanford University, Stanford, California, U.S.A.

Owing to the symmetry with respect to past and present in the definition of the Markov property, this property is preserved if the direction of time is reversed in a process, but the temporal homogeneity is in general not. Now a reversal preserving the latter is of great interest because many analytic and stochastic properties of a process seem to possess an inner duality and deeper insights into its structure are gained if one can trace the paths backwards as well as forwards, as in human history. Such is for instance the case with Brownian motion where the symmetry of the Green's function and the consequent reversibility plays a leading role. Such is also the case of Markov chains where for instance the basic notion of first entrance has an essential counterpart in last exit, a harder but often more powerful tool. Indeed there are many results in the general theory of Markov processes which would be evident from a reverse point of view but are not easy to apprehend directly.

The question of reversal has of course been considered by many authors.(²) One early line of attack (see e.g., [16]) hinged on finding a stationary distribution for the process; once such a distribution is found it is relatively easy to calculate the transition probabilities of the stationary process reversed in time. A more general approach is to reverse the process $\{X_t\}$ from a random time α to get a process $\tilde{X}_t = X_{\alpha-t}$. Hunt [8] considered such a reversal from last exit times in a discrete parameter Markov chain. Chung [4] observed that this could be done with more dispatch from the life time of a continuous parameter minimal chain. Going to a general state space, Ikeda, Nagasawa and Sato [10] considered reversal from the life time of certain processes. This was extended by Nagasawa [15], who reversed more general types of processes from L-times, natural generalizations of last exit times, and later by Kunita and T. Watanabe [11]. An assumption common to

(¹) Research supported in part by the Office of Scientific Research, Office of Aerospace Research, United States Air Force, under AFOSR contract F 44620–67–C–0049, at Stanford University, Stanford, California.

(²) No previous literature on reversal is used in this paper.

these papers is the existence of semigroups or resolvents in duality. Some of the results in this direction have been neatly summarized in [2], [3], [14], [17].

Our approach here is quite different in that, having defined the reverse process $\tilde{X}_t$ as above with α the life time of X_t, we derive the existence of a reverse transition function by showing that the reverse process is indeed a homogeneous Markov process. Our assumptions all bear on the original process, never going beyond those for a Hunt (or standard) process minus the quasi-left continuity. Our fullest result states that we can always reverse such a process from its life time whenever finite to obtain a "moderately strong" homogeneous Markov process, and we give an explicit construction of its transition function. Finally, the restriction to life time will turn out to be only an apparent one, because any such reversal time can be shown to be necessarily the life time of a subprocess. This last important point, requiring compactifications of the state space in its proof, will however not be proved in this paper and will be published later by the second-named author.

Our method takes off from the case of reversal of a minimal Markov chain mentioned earlier (see also [5]). The interesting thing is that this method, which is apparently limited to the special situation of a discrete state space there, can be adapted to the general setting by a natural stretching-out of the life time which renders the smoothness needed for analytic manipulations. The stretching-out is finally removed by probabilistic considerations based on the notion of "essential limit" leading to an "almost fine topology". This notion seems to combine the advantages of separability and shift-invariance and may well turn out to be an essential tool in similar investigations. However, we content ourselves with these remarks here without amplifying them.

1. The finite dimensional distributions of the reverse process

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $(E, \mathcal{E})$ be a locally compact separable metric space and its Borel field. Let $\{X_t, t \geqslant 0\}$ be a homogeneous Markov process with respect to the increasing family of Borel subfields $\{\mathcal{F}_t, t \geqslant 0\}$ of $\mathcal{F}$ and taking values in E; $\mu_t(B)$ and $P_t(x, B)$, $t \geqslant 0$, $x \in E$, $B \in \mathcal{E}$, respectively its absolute distribution and transition function. This means the following:

(i) for each t and x, $B \to P_t(x, B)$ is a probability measure on $\mathcal{E}$;

(ii) for each t and B, $x \to P_t(x, B)$ is in $\mathcal{E}$;

(iii) for each s, t, x and B, we have

$$P_{s+t}(x, B) = \int_E P_s(x, dy)\, P_t(y, B);$$

(iv) for each s, t and B in $\mathcal{E}$, we have with probability one:

$$P\{X_{s+t}\in B\,|\,\mathcal{F}_s\} = P_t(X_s, B).$$

The terminology and notation used above is roughly the same as in [2; p. 14]. Condition (iii), the Chapman–Kolmogorov equation, may be dispensed with; if so we shall qualify the transition function as one "in the loose sense". We need this generalization below. Observe that condition (iv), the Markov property, implies with (i) and (ii) the loose version of (iii) as follows. For each r, s and t, we have with probability one:

$$P_{s+t}(X_r, B) = \int_E P_s(X_r, dy)\, P_t(y, B). \tag{1.1}$$

This often suffices instead of (iii).

Furthermore, we shall assume that the Borel fields $\{\mathcal{F}_t\}$ are augmented with all sets of probability zero. Phrases such as "almost surely" and "for a.e. ω" will mean "for all ω except a set $N\in\mathcal{F}$ with $P(N)=0$". Our first basic hypothesis is that all sample paths of the process X are right continuous. Only later in § 6 will we add the hypothesis that they have also left limits everywhere and finally that X is strongly Markovian. It is of great importance to remember that we are dealing with a fixed process with given initial distribution, and not a family of processes with arbitrary initial values as is customary in Hunt's theory. Thus, convenient notation such as P^x and E^x will not be used.

An "optional time" T is a random variable such that for each t, $\{T<t\}\in\mathcal{F}_t$. The Borel field of sets Λ in $\mathcal{F}$ such that $\Lambda\cap\{T<t\}\in\mathcal{F}_t$ for each t is denoted by $\mathcal{F}_{T+}$. If "$<$" is replaced by "$\leqslant$" in both occurrences above, T will be called "strictly optional" and $\mathcal{F}_{T+}$ replaced by $\mathcal{F}_T$.

Let Δ be an "absorbing state" in E, namely one with the property that if a path ever takes the value Δ it will remain there from then on. There may be more than one such state but one has been singled out. Put

$$\alpha(\omega) = \inf\{t>0: X_t(\omega) = \Delta\},$$

where, as later in all such definitions, the inf is taken to be $+\infty$ when the set in the braces is empty. It is easy to see that α is an optional time, to be called the "life time" of the process. We shall be concerned with reversing the process from such a life time whenever it is finite. Observe that this situation obtains if our process is obtained as a subprocess by "killing" a bigger one in some appropriate manner.

For each x, $t\to P_t(x,\Delta)$ is a distribution function to be denoted by $L(x,t)$. [If the process starting at x were defined, this would be the distribution of its life time.] Our method of reversal relies, *au préalable*, on the following assumptions:

(H1) for each $x \neq \Delta$, $L(x, t)$ is an absolutely continuous function of t with density function $l_t(x)$;

(H2) for each $x \neq \Delta$, $t \to l_t(x)$ is equi-continuous on $(0, \infty)$ with respect to x.

Note that condition (H2) coupled with the fact that $\int_0^\infty l_t(x)\,dt \leqslant 1$ implies that $l_t(x)$ is uniformly bounded on compact subsets of $(0, \infty)$. These conditions seem strong but we shall show later that *both can be entirely removed* if X is assumed to be strongly Markovian (Theorem 6.4). Even without this assumption, their removal will still leave us meaningful and tangible results (Theorem 4.1).

We begin with two lemmas. Throughout the paper we shall use popular concise notation such as $P_s f(x) = \int_E P_s(x, dy) f(y)$.

LEMMA 1.1. *For each $x \neq \Delta$ and $s \geqslant 0, t > 0$ we have*

$$l_{s+t}(x) = P_s l_t(x).$$

Proof. We have if $0 < u < v$,

$$\int_u^v l_{s+r}(x)\,dr = L(x, s+v) - L(x, s+u) = \int_E P_s(x, dy) \int_u^v l_r(y)\,dr = \int_u^v P_s l_r(x)\,dr,$$

where the second equation is a consequence of the Chapman–Kolmogorov equation. It follows that for each fixed $x \neq \Delta$ and $s \geqslant 0$,

$$P_s l_r(x) = l_{s+r}(x)$$

holds for almost all r (Lebesgue measure). Now (H2) implies that both members above are continuous in r, hence the equation holds for all $r > 0$.

LEMMA 1.2. *For each $s > 0$, $t > 0$, and sequence $t_n \downarrow t$, $\lim_{n\to\infty} l_s(X_{t_n})$ exists almost surely; it is equal to $l_s(X_t)$ almost surely provided that for each t, $\mathfrak{F}_t = \mathfrak{F}_{t+}$.*

Proof. Let $0 < t' - t < s$; then by Lemma 1.1,

$$l_s(x) = P_{t'-t}\, l_{s-t'+t}(x).$$

It follows by the Markov property that a.s.

$$l_s(X_{t_n}) = E\{l_{s-t'+t_n}(X_{t'}) \mid \mathfrak{F}_{t_n}\}. \tag{1.2}$$

If $t_n \downarrow t$, then $l_{s-t'+t_n}(X_{t'}) \to l_{s-t'+t}(X_{t'})$ by (H2) and consequently the right member above converges a.s. to $E\{l_{s-t'+t}(X_{t'}) \mid \mathfrak{F}_{t+}\}$. This last step is a case of a useful remark due to Hunt [9], which will be referred to later as Hunt's lemma:

Suppose that the sequence of random variables $\{X_n\}$ converges dominatedly to X_∞ and the sequence of Borel fields $\{\mathfrak{F}_n\}$ is monotone with limit $\mathfrak{F}_\infty$. Then

$$\lim_n E\{X_n \mid \mathfrak{F}_n\} = E\{X_\infty \mid \mathfrak{F}_\infty\}.$$

Remark. Equation (1.2) above remains true even if the transition function of X is in the loose sense, as follows easily from (1.1). Thus Lemma 1.2 remains in force, and condition (iii) may be omitted since it will not be needed again.

The *potential measure* G is defined as follows: for each $A \in \mathcal{E}$:

$$G(A) = \int_0^\infty \mu_t(A)\,dt = E\left\{\int_0^\infty 1_A(X_t)\,dt\right\}.$$

Since the process need not be transient, G may not be a Radon measure. However, we shall presently prove a certain finiteness for it. For each $s>0$, define the measure K_s on $\mathcal{E}$ by

$$K_s(A) = \int_0^\infty \mu_t[1_A l_s]\,dt.$$

We have by Fubini's theorem

$$K_s(E) = \int_0^\infty \mu_0[P_t l_s]\,dt = \mu_0\left[\int_0^\infty l_{t+s}\,dt\right] = \mu_0\left[\int_s^\infty l_t\,dt\right] = P\{s<\alpha<\infty\} \leqslant 1.$$

Hence if $f \in \mathcal{E}$ and f is dominated by l_s for some s, then $Gf<\infty$; in particular G is σ-finite on $\bigcup_{s>0}\{x: l_s(x)>0\}$ and so another application of Fubini's theorem yields

$$K_s(A) = \int_A G(dx)\,l_s(x). \tag{1.3}$$

Now we define the *reverse process* $\tilde{X} = \{\tilde{X}_t, t>0\}$ as follows. Adjoin a new point $\tilde{\Delta}$ to E, where $\tilde{\Delta} \notin E$ and $\tilde{\Delta}$ is isolated in $E \cup \tilde{\Delta}$: put

$$\tilde{X}_t = \begin{cases} X_{\alpha-t} & \text{if } 0<t\leqslant\alpha<\infty; \\ \tilde{\Delta} & \text{if } \alpha<\infty, t>\alpha; \\ \tilde{\Delta} & \text{if } \alpha=\infty, t>0. \end{cases} \tag{1.4}$$

The sample paths of $\tilde{X}$ are therefore just those of X with t reversed in direction, apart from trivial completions; hence they are left continuous. Furthermore, $\tilde{X}$ never takes the value Δ and it takes the value $\tilde{\Delta}$ wherever it is not in E. Hence when we specify its absolute distributions and transition probabilities we may confine ourselves to subsets of E, as we do in the theorem below.

THEOREM 1.1. *Under hypotheses* (H1) *and* (H2), *the absolute distribution of* $\tilde{X}_s$ *is* K_s *given by* (1.3), *the joint distribution of* $\tilde{X}_s$ *and* $\tilde{X}_t$, $0 < s \leqslant t$, *is given by*

$$P\{\tilde{X}_s \in A, \tilde{X}_t \in B\} = \int_B G(dx) \int_A P_{t-s}(x, dy)\, l_s(y), \tag{1.5}$$

where $A \in \mathcal{E}$, $B \in \mathcal{E}$. *More generally, if* $0 < t_1 < t_2 < \ldots < t_n$ *and* $A_1, \ldots, A_n$ *all belong to* $\mathcal{E}$ *we have*

$$P\{\tilde{X}_{t_j} \in A_j, 1 \leqslant j \leqslant n\} = \int_{A_n} G(dx_n) \int_{A_{n-1}} P_{t_n - t_{n-1}}(x_n, dx_{n-1}) \ldots \int_{A_1} P_{t_2 - t_1}(x_2, dx_1)\, l_{t_1}(x_1). \tag{1.6}$$

Note: As remarked following Lemma 1.2, the P_t's may be transition functions in the loose sense.

Proof. We shall prove only (1.5) which contains the main argument; the proof of (1.6) requires no new argument while the assertion about absolute distributions follows from (1.5) if we take $t = s$ and $B = A$ there.

Let C_K^+ denote the class of positive continuous functions on $E \cup \tilde{\Delta}$ with compact support and vanishing at Δ and $\tilde{\Delta}$. It is sufficient to prove that for each f and g in C_K^+, we have

$$E\{f(\tilde{X}_s)\, g(\tilde{X}_t)\} = G[gP_{t-s}(f l_s)]. \tag{1.7}$$

We do this by a discrete approximation. Set

$$\alpha_n = [2^n \alpha + 1]\, 2^{-n},$$

where $[2^n \alpha + 1]$ is the greatest integer $\leqslant 2^n \alpha + 1$; then $\alpha_n > \alpha$ and $\alpha_n \downarrow \alpha$ as $n \to \infty$. Since X has right continuous paths, the left number of (1.7) is equal to the limit of

$$E\{f \circ X(\alpha_n - s) \cdot g \circ X(\alpha_n - t)\} \tag{1.8}$$

as $n \to \infty$, where "$\circ$" denotes composition of functions. For each integer N, write (1.8) as

$$\sum_{2^n t \leqslant k \leqslant 2^n N} E\{f \circ X(k2^{-n} - s) \cdot g \circ X(k2^{-n} - t); \alpha_n = k2^{-n}\}$$
$$+ E\{f \circ X(\alpha_n - s) \cdot g \circ X(\alpha_n - t); N < \alpha < \infty\}. \tag{1.9}$$

The second term tends to zero as $N \to \infty$ uniformly in n. Observing that

$$E\{\alpha_n = k2^{-n} \mid \mathcal{F}_{k2^{-n}-s}\} = \int_{s-2^{-n}}^{s} l_r \circ X(k2^{-n} - s)\, dr,$$

we write the kth term of the sum in (1.0) as

$$E\{f\circ X(k2^{-n}-s)\cdot g\circ X(k2^{-n}-t)\cdot\int_{s-2^{-n}}^{s} l_r\circ X(k2^{-n}-s)\,dr\}$$

$$=\int_E \mu_{k2^{-n}-t}(dx)\,g(x)\int_E P_{t-s}(x,dy)\,f(y)\,\{2^{-n}l_s(y)+\int_{s-2^{-n}}^{s}[l_r(y)-l_s(y)]\,dr\}$$

$$=2^{-n}\mu_{k2^{-n}-t}[gP_{t-s}(fl_s)]+F_{kn},$$

where

$$F_{kn}=\int_E \mu_{k2^{-n}-t}(dx)\,g(x)\int_E P_{t-s}(x,dy)\,f(y)\int_{s-2^{-n}}^{s}[l_r(y)-l_s(y)]\,dr.$$

We have,

$$\sum_{2^nt\leqslant k\leqslant 2^nN} F_{kn}\leqslant 2^n(N-t)\,\|f\|\,\|g\|2^{-n}\sup_{y\in E}\sup_{|r-s|\leqslant 2^{-n}}|l_r(y)-l_s(y)|$$

which converges to zero as $n\to\infty$ by (H2), for each N. It remains to evaluate the limit as $n\to\infty$ of

$$\sum_{2^nt\leqslant k\leqslant 2^nN} 2^{-n}\mu_{k2^{-n}-t}[gP_{t-s}(fl_s)]. \tag{1.10}$$

Consider the function

$$u\to\mu_u[gP_{t-s}(fl_s)]=E\{g(X_u)\,f(X_{u+t-s})\,l_s(X_{u+t-s})\}. \tag{1.11}$$

Clearly $u\to g(X_u)f(X_{u+t-s})$ is right continuous. Since $l_s(x)$ is bounded in x by (H2), it follows from Lemma 1.2 and Lebesgue's bounded convergence theorem that the function in (1.11) has a right limit everywhere. Hence it is integrable in the Riemann sense and consequently the limit of (1.10) is the Riemann (ergo Lebesgue) integral

$$\int_0^N \mu_u[gP_{t-s}(fl_s)]\,du.$$

Letting $N\to\infty$ we obtain the right member of (1.7), which is finite by the remarks preceding (1.3), q.e.d.

2. The transition function of the reverse process

We prove in this section that the reverse process is temporally homogeneous and exhibit a loose transition function $\tilde{P}_t(y,A)$ for it. If such a function exists, it must be the Radon–Nikodym derivative

$$\frac{P\{\tilde{X}_s\in dy,\tilde{X}_{s+t}\in A\}}{P\{\tilde{X}_s\in dy\}}.$$

The problem is to define this measurably in y and simultaneously for all A in $\mathcal{E}$. Doob [6] has given a similar procedure in connection with conditional probability distributions in

the wide sense which has been extended by Blackwell [1] to a more general space. We shall indicate a simpler argument using the functional approach.

Define the function h on E by

$$h(x)=\int_0^\infty e^{-s}l_s(x)\,ds.$$

We have by Lemma 1.1,

$$P_th=\int_0^\infty e^{-s}P_tl_s\,ds=e^t\int_t^\infty e^{-s}l_s\,ds\leqslant e^th,$$

from which it follows that h is 1-excessive with respect to (P_t). Furthermore, $h(x)=0$ if and only if $l_s(x)=0$ for all s by the continuity of $s\to l_s(x)$. Next, recalling (1.3), we define the measure K on $\mathcal{E}$ by

$$K(A)=\int_0^\infty e^{-s}K_s(A)\,ds=\int_A G(dx)\,h(x). \tag{2.1}$$

Since $K_s(E)\leqslant 1$ for each s we have $K(E)\leqslant 1$.

Now for each t we define a function Π_t on product Borel sets of $E\times E$ by

$$\Pi_t(A,B)=\int_A G(dx)\int_B P_t(x,dy)\,h(y). \tag{2.2}$$

It follows from (2.1) that

$$\Pi_t(A,E)\leqslant\int_A G(dx)\,e^th(x)\leqslant e^tK(A); \tag{2.3}$$

on the other hand, since $GP_t\leqslant G$, we have

$$\Pi_t(E,B)\leqslant\int_B G(dy)\,h(y)=K(B). \tag{2.4}$$

Consequently $\Pi_t(A,\cdot)$ and $\Pi_t(\cdot,B)$ are both measures which are absolutely continuous with respect to K $(<<K)$.

THEOREM 2.1. *The reverse process $\{\tilde{X}_t, t>0\}$ is a homogeneous Markov process taking values in $E\cup\tilde{\Delta}$, with a version of the Radon–Nikodym derivative*

$$\frac{\Pi_t(A,dy)}{K(dy)}=\tilde{P}_t(y,A),\quad t\geqslant 0,$$

as its transition function in the loose sense.

Note: $\tilde{P}_0(y,A)=\varepsilon_A(y)$.

Proof. Let D_0 be a countable dense subset of C_K. Let D be the smallest class of functions on E containing D_0 and the constant 1 which is closed under addition and multiplication by a rational number. D is countable and contains all rational constants. Fix $t>0$; for each f in D and B in $\mathcal{E}$, we put

$$\Pi_t(f, B) = \int_B f(x)\, \Pi_t(dx, B).$$

As a signed measure $\Pi_t(f, \cdot) << K(\cdot)$ by (2.4). Let $L(f, y)$ denote a version of the Radon–Nikodym derivative $\Pi_t(f, dy)/K(dy)$ such that $y \to L(f, y)$ is in $\mathcal{E}$ for each $f \in D$. There is a set Z in $\mathcal{E}$ with $K(Z)=0$, such that if $y \in E-Z$, then

(a) $L(f, y) \geqslant 0$ if $f \in D$, $f \geqslant 0$;
(b) $L(0, y)=0$;
(c) $L(cf, y) = cL(f, y)$ if $f \in D$ and $cf \in D$ where c is real;
(d) $L(f+g, y) = L(f, y) + L(g, y)$ if $f \in D$, $g \in D$;
(e) $|L(f, y)| \leqslant \|f\|$ if $f \in D$.

The proofs of these assertions are all trivial. E.g., to show (c), we write for each $B \in \mathcal{E}$,

$$\int_B L(cf, y)\, K(dy) = \Pi_t(cf, B) = c\Pi_t(f, B) = \int_B cL(f, y)\, K(dy)$$

and take B to be $\{y: L(cf, y) > cL(f, y)\}$ or $\{y: L(cf, y) < cL(f, y)\}$. It follows that the relation in (c) holds for each pair f and cf in D, for K-a.e. y. Since D is countable, this establishes (c).

For $y \in E-Z$, and $f \in C_K$, we define

$$L(f, y) = \lim_n L(f_n, y), \tag{2.5}$$

where $\{f_n\}$ is any sequence in D which converges to f in norm. It follows from (e) that the limit above exists and does not depend on the choice of the sequence. It is trivial to verify that $L(\cdot, y)$ so extended to C_K is a linear functional over the real coefficient field with norm $\leqslant 1$. To see that it is positive, let $f \in C_K$, $f \geqslant 0$; then for every rational $\varepsilon > 0$, we have $f+\varepsilon \geqslant \varepsilon$. Hence if $\|f_n - f\| \to 0$, then $f_n + \varepsilon \geqslant 0$ for sufficiently large n. It follows from (a) above and (2.5) that $L(f+\varepsilon, y) \geqslant 0$, and hence, by linearity and (c), that $L(f, y) \geqslant 0$.

Thus the linear functional $L(\cdot, y)$ defined in (2.5) is a Radon measure on $\mathcal{E}$ with total mass $\leqslant 1$. We now put for $y \in E-Z$ and $A \in \mathcal{E}$:

$$\hat{P}_t(y, A) = L(A, y),$$

$$\hat{P}_t(y, \tilde{\Delta}) = 1 - L(E, y);$$

for $y \in Z$ and $A \in \mathcal{E}$: $\qquad \hat{P}_t(y, A) = \varepsilon_{\{y\}}(A)$;

finally $\qquad \hat{P}_t(\tilde{\Delta}, \{\tilde{\Delta}\}) = 1$.

Then for $y \in E \cup \{\tilde{\Delta}\}$, $A \to \hat{P}_t(y, A)$ is a probability measure on $\mathcal{E}_{\tilde{\Delta}}$, the Borel field generated by $\mathcal{E}$ and $\tilde{\Delta}$; $y \to \hat{P}_t(y, A)$ is in $\mathcal{E}_{\tilde{\Delta}}$ for each $A \in \mathcal{E}$; and we have

$$\Pi_t(A, B) = \int_B \hat{P}_t(y, A)\, K(dy). \tag{2.6}$$

Thus $\hat{P}_t(y, A)$ will be a transition function in the loose sense for $\tilde{X}$ provided we can verify the relation corresponding to (iv) at the beginning of § 1, namely that we have with probability one:

$$P\{\tilde{X}_{s+t} \in A \mid \tilde{\mathfrak{F}}_s\} = \hat{P}_t(\tilde{X}_s, A); \tag{2.7}$$

where for each $t > 0$, $\tilde{\mathfrak{F}}_s$ is the Borel field generated by $\{\tilde{X}_r, 0 < r \leqslant s\}$ and augmented with all sets of probability zero.

Equivalently, we may verify that the finite-dimensional distributions of $\tilde{X}$ obtained in Theorem 1.1 can be written in the proper form by means of K_s and $\hat{P}_t$ as shown below. We begin with the following lemma which embodies the ***duality relation*** mentioned in the introduction.

LEMMA 2.1. *For every positive g in $\mathcal{E} \times \mathcal{E}$ such that g vanishes on the set $E \times \{y: h(y) = 0\}$, we have*

$$\int_E G(dx) \int_E P_t(x, dy)\, g(x, y) = \int_E G(dy) \int_E \hat{P}_t(y, dx)\, g(x, y). \tag{2.8}$$

Proof. If $g(x, y) = 1_A(x)\, 1_B(y)\, h(y)$, $A \in \mathcal{E}$, $B \in \mathcal{E}$, then the left member of (2.8) is just

$$\int_A G(dx) \int_B P_t(x, dy)\, h(y) = \Pi_t(A, B) = \int_B \hat{P}_t(y, A)\, K(dy)$$

by (2.6), which reduces to

$$\int_B \hat{P}_t(y, A)\, h(y)\, G(dy) = \int_E G(dy) \int_E \hat{P}_t(y, dx)\, 1_A(x)\, 1_B(y)\, h(y).$$

Hence (2.8) is true for g of the specified form, and so is true for all positive g of the form fh, where $f \in \mathcal{E} \times \mathcal{E}$, by a familiar monotone class argument. Now it is trivial that this coincides with the class of g stated in the lemma, q.e.d.

Returning to the proof of Theorem 2.1, let us define for $0 < t_1 < \ldots < t_n$ and arbitrary $x_1, \ldots, x_n$ in E:

$$k_1(x_1) = l_{t_1}(x_1),$$

$$k_n(x_n) = \int_{A_{n-1}} P_{t_n-t_{n-1}}(x_n, dx_{n-1}) \dots \int_{A_1} P_{t_2-t_1}(x_2, dx_1)\, l_{t_1}(x_1), \quad n \geq 2.$$

It follows from Lemma 1.1 that

$$k_n(x_n) \leq P_{t_n-t_{n-1}} \dots P_{t_2-t_1} l_{t_1}(x_n) = l_{t_n}(x_n)$$

so that k_n vanishes where h does. We have therefore by repeated application of Lemma 2.1 to the right member of (1.6):

$$\int_{A_n} G(dx_n) \int_{A_{n-1}} P_{t_n-t_{n-1}}(x_n, dx_{n-1})\, k_{n-1}(x_{n-1})$$

$$= \int_{A_{n-1}} G(dx_{n-1})\, k_{n-1}(x_{n-1})\, \hat{P}_{t_n-t_{n-1}}(x_{n-1}, A_n)$$

$$= \int_{A_{n-1}} G(dx_{n-1}) \int_{A_{n-2}} P_{t_{n-1}-t_{n-2}}(x_{n-1}, dx_{n-2})\, k_{n-2}(x_{n-2})\, \hat{P}_{t_n-t_{n-1}}(x_{n-1}, A_n)$$

$$= \int_{A_{n-2}} G(dx_{n-2})\, k_{n-2}(x_{n-2}) \int_{A_{n-1}} \hat{P}_{t_{n-1}-t_{n-2}}(x_{n-2}, dx_{n-1})\, \hat{P}_{t_n-t_{n-1}}(x_{n-1}, A_n)$$

$$= \dots = \int_{A_1} G(dx_1)\, l_{t_1}(x_1) \int_{A_2} \hat{P}_{t_2-t_1}(x_1, dx_2)$$

$$\dots \int_{A_{n-1}} \hat{P}_{t_{n-1}-t_{n-2}}(x_{n-2}, dx_{n-1})\, \hat{P}_{t_n-t_{n-1}}(x_{n-1}, A_n).$$

Comparing this with Theorem 1.1 we see that $\tilde{X}$ has indeed $\hat{P}_t$ as a version of its transition function and Theorem 2.1 is completely proved.

3. A regularity property of the reverse transition function

We shall show that an arbitrary collection of versions of the Radon–Nikodym derivatives $\{\hat{P}_t, t>0\}$ obtained in Theorem 2.1 has certain regularity properties and use these to construct a "standard modification" that is *vaguely left continuous* in t. This results from the fact that $\hat{P}_t(y, A)$ is the loose-sense transition function of a homogeneous Markov process whose sample paths are all left continuous, and will be stated in this general form, using the notation X_t and P_t instead of $\tilde{X}_t$ and $\hat{P}_t$.

From now on we write R for $[0, \infty)$ and Q for an arbitrary countable dense subset of R. To alleviate the notation *we shall reserve in this section the letters r and r' to denote members of Q. Thus, for instance, $r \to t$ means $r \in Q$ and $r \to t$.* The notation $s \to t+$ means

$s > t$ and $s \to t$, similarly $s \to t-$ means $s < t$ and $s \to t$. If X_t is a Markov process with absolute distributions μ_t, a set Z in the completion of $\mathcal{E}$ with respect to all $\{\mu_t,\ t > 0\}$ such that $\mu_t(Z) = 0$ for all $t > 0$ will be called "insignificant".

THEOREM 3.1. *Suppose $\{X_t, t > 0\}$ is a homogeneous Markov process taking values in E and having left continuous sample paths. Suppose $P_t(x, B)$ is its transition function in the loose sense and $\mu_t(B)$ its absolute distribution. Then the following two assertions are true.*

(a) *For each Q there is an insignificant set Z such that for every $x \notin Z$ and $f \in C_K$, we have*

$$\forall t > 0: \lim_{r \to t-} P_r f(x) \quad \textit{exists.}$$

(b) *For each $t > 0$, there is an insignificant set Z_t such that for every $x \notin Z_t$ and $f \in C_K$, we have*

$$\lim_{r \to t-} P_r f(x) = P_t f(x).$$

Remark. There is an obvious analogue if X has right continuous paths.

Proof of (a). Let $\varepsilon > 0$, $f \in C_K$ and put

$$H = \{(t, x): \varliminf_{s \to t-} P_s f(x) < \varlimsup_{s \to t-} P_s f(x) - \varepsilon\}. \tag{3.1}$$

If Π denotes the projection of $R \times E$, we have

$$\Pi(H) = \{x: \exists t > 0: \varliminf_{s \to t-} P_s f(x) < \varlimsup_{s \to t-} P_s f(x) - \varepsilon\}. \tag{3.2}$$

To prove (a) it is sufficient to show that $\Pi(H)$ is insignificant and this will be done by a capacity argument due to P. A. Meyer [12]. We sketch the set-up below; note that a "$\mathcal{k}$-analytic" set below is an "analytic" or "Souslin" set in the classical sense.

Let $\mathcal{B}$ be the Borel field of R, $\mathcal{C}$ the class of compact sets of R, $\mathcal{k}$ the class of compact sets of E. It is easy to see that $H \in \mathcal{B} \times \mathcal{E}$ because Q is countable (cf. e.g., [5; pp. 161–2]), hence $\Pi(H)$ is $\mathcal{k}$-analytic and so measurable with respect to the completed measure μ_s.

LEMMA 3.1. *If $s > 0$, there exists $L \in \mathcal{E}$, such that $L \subset \Pi(H)$ with $\mu_s(L) = \mu_s(\Pi(H))$, and a strictly positive $\mathcal{E}$-measurable function τ defined on L whose graph*

$$\{(x, \tau(x)): x \in L\}$$

is contained in H.

This is a particular case of Meyer's theorem but can be proved quickly as follows. For every subset A of $R \times E$ define

$$\varphi(H) = \mu_s^*(\Pi(H)),$$

where μ_s^* is the outer measure induced by μ_s. Then φ is a capacity and H is analytic, both with respect to the class of compact sets of the product space $R \times E$. Hence H is φ-capacitable and there is a compact subset K_1 of H such that $\varphi(K_1) > \varphi(H)/2$. Now define τ_1 on $L_1 = \Pi(K_1)$ by

$$\tau_1(x) = \inf \{t: (t, x) \in K_1\}.$$

The compactness of K_1 implies, first that $(x, \tau_1(x)) \in K_1$ and second that for each real c, the set $\{x: \tau_1(x) \leqslant c\}$ is closed so that τ_1 is $\mathcal{E}$-measurable, indeed lower semi-continuous. [We owe Professor Wendell Fleming the last remark which replaces a longer argument.] If we choose an increasing sequence of compact K_n with $\varphi(K_n) \uparrow \varphi(H)$ and define the corresponding L_n and τ_n as above we see that the set $L = \bigcup_n L_n$ and the function τ such that $\tau(x) = \tau_n(x)$ for $x \in L_n - L_{n-1}$ (with $L_0 = \varnothing$) satisfy the requirements, q.e.d.

It follows from the lemma that for every $x \in L$, we have

$$\varliminf_{r \to \tau(x)-} P_r f(x) < \varlimsup_{r \to \tau(x)-} P_r f(x) - \varepsilon. \tag{3.3}$$

Hence if we define two subsets of Q as follows:

$$\Gamma_1(x) = \left\{ r \in Q : P_r f(x) > \varlimsup_{r' \to \tau(x)-} P_{r'} f(x) - \frac{\varepsilon}{3} \right\},$$

$$\Gamma_2(x) = \left\{ r \in Q : P_r f(x) < \varliminf_{r' \to \tau(x)-} P_{r'} f(x) + \frac{\varepsilon}{3} \right\};$$

then for every $x \in L$, $\tau(x)$ is an accumulation point from the left of both $\Gamma_1(x)$ and $\Gamma_2(x)$, namely that for every $\delta > 0$, we have $(\tau(x) - \delta, \tau(x)) \cap \Gamma_i(x) \neq \varnothing$, $i = 1, 2$. It follows from this that for either i we can construct $\mathcal{E}$-measurable functions σ_n on L, taking values in $\Gamma_i(x)$, and such that $\sigma_n(x) \to \tau(x)-$ for all x in L. This is a familiar construction of which a more elaborate form will be stated in § 6. Assuming this, we are ready to prove (a). Let $\{\sigma_n'\}$ and $\{\sigma_n''\}$ be the $\{\sigma_n\}$ just mentioned corresponding to Γ_1 and Γ_2 respectively, and let $\{\tau_n\}$ be the alternating sequence $\{\sigma_1', \sigma_1'', \sigma_2', \sigma_2'', \ldots\}$. We have then for every $x \in L$:

$$P_{\sigma_n'(x)} f(x) > P_{\sigma_n''(x)} f(x) - \frac{\varepsilon}{3}. \tag{3.4}$$

Now consider the equation

$$\int_L \mu_s(dx) P_{\tau_n(x)} f(x) = E\{X_s \in L;\ f \circ X(s + \tau_n(X_s))\}, \tag{3.5}$$

which is a consequence of the Markov property of X since τ_n is countably-valued. The

right member of (3.5) converges as $n\to\infty$ to the limit obtained by replacing τ_n with τ there, since X has left continuous paths. But by (3.4) the left member cannot converge unless $\mu_s(L)=0$. This must then be true and so $\mu_s(\Pi(H))=0$ by Lemma 3.1. Since s is arbitrary, $\Pi(H)$ is insignificant. Writing $H_f(\varepsilon)$ for this H, letting f run through a countable set D dense in C_K, and setting $Z=\bigcup_{f\in D}\bigcup_{n=1}^{\infty}H_f(n^{-1})$, we obtain (a).

The proof of (b) is similar but simpler. For a fixed $t>0$, consider

$$H_t=\left\{x:\varliminf_{r\to t-} P_rf(x)<P_tf(x)-\varepsilon\right\}.$$

$$\Gamma_t(x)=\left\{r\in Q: P_rf(x)<P_tf(x)-\frac{\varepsilon}{2}\right\}.$$

Then $H_t\in\mathcal{E}$ (no capacity argument is needed here), and there exist $\mathcal{E}$-measurable functions τ_n defined on H_t, taking values in $\Gamma_t(x)$, and increasing to t as $n\to\infty$. It follows that

$$\int_{H_t}\mu_s(dx)\left(P_tf(x)-\frac{\varepsilon}{2}\right)\geqslant\int_{H_t}\mu_s(dx)P_{\tau_n(x)}f(x)$$

$$=E\{X_s\in H_t; f\circ X(s+\tau_n(X_s))\}\to E\{X_s\in H_t; f\circ X(s+t)\}=\int_{H_t}\mu_s(dx)\,P_tf(x).$$

Hence $\mu_s(H_t)=0$. Together with a symmetric argument on the upper limit, this establishes (b).

THEOREM 3.2. *Under the hypotheses of Theorem 3.1, there exists a transition function $P_t^*(x,B)$ in the loose sense for the process X such that for each $f\in C_K$ we have*

$$\forall t>0;\ \lim_{s\to t-} P_s^*f=P_t^*f.$$

This means: for each x, $t\to P_t^*(x,\cdot)$ is vaguely left continuous as measures. We shall write a vague limit in this sense as "v lim" below.

Proof. In view of (a) of the preceding theorem, we may define

$$\forall t>0, x\notin Z:\quad P_t^*(x,\cdot)=\mathrm{v}\lim_{r\to t-} P_r(x,\cdot)$$

$$\forall t>0, x\in Z:\quad P_t^*(x,\cdot)=\delta_x(\cdot)=P_0^*(x,\cdot).$$

For each $f\in C_K$, $x\to P_t^*f(x)$ is in $\mathcal{E}$. By (b) of the theorem, we have for every s almost surely

$$P_t^*(X_s,f)=P_t(X_s,f),$$

and consequently P_t^* as well as P_t serves as a transition function in the loose sense. Finally, from $P_t^* f = \lim_{r \to t-} P_r f$ ($r \in Q$!) it follows that $t \to P_t^* f$ is left continuous, q.e.d.

Applying Theorem 3.2 to the reverse process $\tilde{X}$ in Theorem 2.1, we conclude that its transition function $\tilde{P}_t(y, \cdot)$ may be modified to be vaguely left continuous in t for each y, as defined above.

4. The removal of assumptions (H1) and (H2)

The preceding results were proved under (H1) and (H2). If the life time α of the given process does not satisfy these conditions, it will now be shown in what sense the results may be carried over. Roughly speaking, they remain true provided that "reversed time" be liberally interpreted as beginning at some fictitious (but by no means nebulous) origin. Or else if this is not allowed, then the results are still true provided that an exceptional set of reversed time of zero Lebesgue measure be ignored. Finally we shall show in § 6 that all fiction or exception may be dropped provided that the given (forward) process is assumed to be strongly Markovian instead of merely Markovian as we do now. This however lies deeper.

For the present a little device suffices: one simply extends the life time from α to α^* by adding exponentially distributed holding times until the distribution of α^*, being the convolution of that of α with smooth densities, achieves the kind of good behavior required by (H1) and (H2). In fact, this device will make the distribution of α^* as smooth as one may wish as a function of t, but only mildly so as a function of x. This will be sufficient since we need only a certain uniformity with respect to x in (H2). Now we can reverse the prolonged process from the new life time α^* by the preceding theorems. The true reversal from α will then appear as the portion of the reversed prolonged process starting from $\alpha^* - \alpha$, which is an optional time for it. Hence if the last-mentioned process is moderately strongly Markovian — as we shall prove in § 6 — the true reverse process will behave in like fashion.

Let β_i, $i = 1, 2, 3$, be three random variables independent of one another and of the Borel field generated by $\{X_t, t \geq 0\}$, and having the common distribution with density $\lambda e^{-\lambda t} dt$, $\lambda > 0$. Adjoin three distinct new points Δ_i, $i = 1, 2, 3$ to E and define the prolonged process as follows:

$$Y_t = \begin{cases} X_t & \text{if } t < \alpha, \\ \Delta_1 & \text{if } \alpha \leq t < \alpha + \beta_1, \\ \Delta_2 & \text{if } \alpha + \beta_1 \leq t < \alpha + \beta_1 + \beta_2, \\ \Delta_3 & \text{if } \alpha + \beta_1 + \beta_2 \leq t < \alpha + \beta_1 + \beta_2 + \beta_3, \\ \Delta & \text{if } t \geq \alpha + \beta_1 + \beta_2 + \beta_3. \end{cases}$$

We shall regard λ as fixed, put

$$\beta=\beta_1+\beta_2+\beta_3, \qquad \alpha^*=\alpha+\beta,$$

and denote the density of β by

$$b(t)=2^{-1}\lambda^3 t^2 e^{-\lambda t}, \quad t\geq 0.$$

Thus the distribution of α^* is given by

$$L^*(x,t)=\int_0^t L(x,s)\,b(t-s)\,ds$$

with the density

$$l^*(x,t)=\int_0^t L(x,s)\,b'(t-s)\,ds.$$

Since

$$|l^*(x,t)|\leq\int_0^t |b'(s)|\,ds\leq\int_0^\infty 2^{-1}\lambda^3|2t-t^2|e^{-\lambda t}\,dt<\infty,$$

and

$$\left|\frac{d}{dt}l^*(x,t)\right|\leq\int_0^t |b''(s)|\,ds\leq\int_0^\infty 2^{-1}\lambda^3|2-4\lambda t+\lambda^2 t^2|e^{-\lambda t}dt<\infty,$$

$l^*(x,t)$ is uniformly bounded in all x and t and $t\to l^*(x,t)$ has a derivative bounded uniformly in x and t and hence is equi-continuous in t with respect to all x. This means (H1) and (H2) hold.

The parameter λ will play no role below, but let us remark that as $\lambda\to\infty$, $Y_t\to X_t$ for all t almost surely. Now we define the reverse process to Y from α^* just as we did the reverse to X from α in (1.4):

$$\tilde{Y}_t=\begin{cases} Y_{\alpha^*-t} & \text{if } 0<t\leq\alpha^*<\infty,\\ \tilde{\Delta} & \text{if } \alpha^*<\infty, t>\alpha^*,\\ \tilde{\Delta} & \text{if } \alpha^*=\infty, t>0.\end{cases}$$

Then $\tilde{Y}_t=\tilde{X}_{t-\beta}$ if $t>\beta$. Since Y satisfies (H1) and (H2), $\tilde{Y}$ satisfies the conclusions of Theorems 1.1 and 2.1.

The independence of $\alpha^*-\alpha$ and $\{X_t, t\geq 0\}$ should be formalized by considering the product measure space $(\Omega\times R, \mathfrak{F}\times\mathcal{B}, P\times\nu)$ where ν is the measure with density b on R. If we regard the Y process as defined on this space and write $\hat{\omega}=(\omega,\omega')$, $Y(t,\hat{\omega})=X(t,\omega)$ if $t<\alpha(\omega')$, etc., then the following lemma is not only obvious but even true (it may be false otherwise).

Lemma 4.1. *Let f_j, $1\leq j\leq n$, be bounded, $\mathcal{E}$-measurable functions vanishing at Δ_i, $i=1,2,3$. Then for each $t_0<t_1$:*

$$E\left\{\beta<t_0;\prod_{j=1}^n f_j(\tilde{Y}_{t_j})\right\}=\int_0^{t_0} E\left\{\prod_{j=1}^n f_j(\tilde{X}_{t_j-s})\right\}b(s)\,ds. \tag{4.1}$$

We now state and prove the result accruing from Theorem 2.1 after the removal of (H1) and (H2).

THEOREM 4.1. *Let $\{X_t, t\geq 0\}$ be a homogeneous Markov process with right continuous paths and life time α. Let $\{\tilde{X}_t, t>0\}$ be the reverse process defined by* (1.1), *and $\tilde{\mu}_t(A)=P\{\tilde{X}_t\in A\}$ for $A\in\mathcal{E}$. Then there exists $\hat{P}_t(x,A)$, $t\geq 0$, $x\in E$, $A\in\mathcal{E}$, satisfying conditions* (i) *and* (ii) *for a transition function given at the beginning of § 1, such that $t\to\hat{P}_t(x,\cdot)$ is vaguely left continuous for each x, with the following property. Given $0\leq t_1<\ldots<t_n$ and $A_0, A_1, \ldots, A_n$ in $\mathcal{E}$, we have for almost every* (*Lebesgue*) *t_0 in* $(0, t_1)$:

$$P\{\tilde{X}_{t_j}\in A_j, 0\leq j\leq n\}=\int_{A_0}\tilde{\mu}_{t_0}(dx_0)\int_{A_1}\hat{P}_{t_1-t_0}(x_0,dx_1)\ldots\int_{A_n}\hat{P}_{t_n-t_{n-1}}(x_{n-1},dx_n). \tag{4.2}$$

Remark. The proof will show how to calculate $\tilde{\mu}_t$ and $\hat{P}_t$.

Proof. Let f_j be as in Lemma 4.1. Since $\tilde{Y}$ satisfies the conclusions of Theorems 2.1 and 3.1, let P_t^* be its transition function in the loose sense having the stated regularity property. We have then for $t<t_0$:

$$E\left\{\beta<t;\prod_{j=1}^{n}f_j(\tilde{Y}_{t_j})\right\}=E\{\beta<t;(f_0\varphi)(\tilde{Y}_{t_0})\}, \tag{4.3}$$

where

$$\varphi(x)=\int_E P^*_{t_1-t_0}(x,dx_1)f_1(x_1)\int_E P^*_{t_2-t_1}(x_1,dx_2)f_2(x_2)\ldots\int_E P^*_{t_n-t_{n-1}}(x_{n-1},dx_n)f_n(x_n).$$

Using Lemma 3.1 in both members of (4.3), we obtain

$$\int_0^t ds\,b(s)E\left\{\prod_{j=0}^{n}f_j(\tilde{X}_{t_j-s})\right\}=\int_0^t ds\,b(s)\,E\left\{(f_0\varphi)(\tilde{X}_{t_0-s})\right\}. \tag{4.4}$$

This being true for all $t<t_0$, and $b(s)>0$ for $s>0$, we conclude that the two expectations in (4.4) are equal for almost all $s<t_0$. Since t_0 is arbitrary, it follows that given $t_1<\ldots<t_n$, we have

$$E\left\{f_0(\tilde{X}_t)\prod_{j=1}^{n}f_j(\tilde{X}_{t_j})\right\}=E\{(f_0\varphi)(\tilde{X}_t)\}$$

for almost all $t<t_1$. This implies (4.2). Note that the loose-sense transition for $\tilde{X}$ may be taken to be that of $\tilde{Y}$ for any $\lambda>0$, and that its absolute distribution $\tilde{\mu}_t$ is determined by the equation below, valid for $t>0$, $A\in\mathcal{E}$;

$$\int_0^t \mu_{t-s}(A)\,b(s)\,ds=P\{\tilde{Y}_t\in A\}=\int_A G(dx)\,l^*(x,t).$$

We end this section with some examples to illustrate the possibilities and limitations:

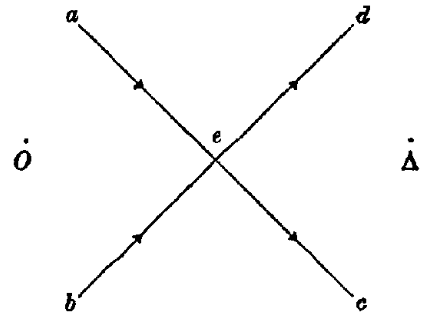


Fig. 1.

Example 1. The state space consists of the two diagonals $\overline{ac}$ and $\overline{bd}$ of a square with side length $\sqrt{2}$ and center e, together with two outside points O and Δ. The process starts at O which is a holding point (with density $e^{-t}dt$ for the holding time distribution), then jumps to a or b with probability $\frac{1}{2}$ each. From either point it moves with unit speed along the diagonal until it reaches c or d, and then jumps to the absorbing point Δ. This process is Markovian but not strongly so, as the strong Markov property fails at the hitting time of e. The reverse process is not Markovian at $t=1$, when it is at the state e. Observe that the transition probabilities for the forward process do not satisfy the Chapman–Kolmogorov equation $P_2(x, d)=P_1(x, e)P_1(e, d)$ for both $x=a$ and $x=b$, no matter how $P_1(e, d)$ is defined.

Example 2. This is an elaboration of the preceding example, in which the reverse process is not Markovian at an uncountable set of t (but of measure 0 in accordance with Theorem 4.1). Let f be a nonnegative continuous function on $[0, 1]$, whose set of zeros is the Cantor set. The state space consists of the graphs of f and of $-f$. The process starts at $(0, 0)$ which is a holding point, then follows either the graph of f or the graph of $-f$ with probability $\frac{1}{2}$ each until it reaches $(0, 1)$ which is the absorbing point. This process is Markovian but not strongly so, and the reverse process is not Markovian, for the Markov property fails at all t in the Cantor set.

Example 3. This example shows that even if the forward process is strongly Markovian, the reverse one need not be so. Let the process be the uniform motion on the line starting at -1, moving to the right until it hits O which is a holding point, after which it jumps to Δ. The reverse process is Markovian but not strongly so, since it has continuous paths and yet starts at a holding point.

5. Essential limits

Let $R=[0, \infty)$ and let "measure" below be the Lebesgue measure on R, denoted by m. For an extended real-valued function f on R, we say that "its essential supremum on a

measurable set S exceeds c" iff there is a subset of S of strictly positive measure on which $f > c$; the supremum of all such c is the ess sup of f on S, unless the set of c is empty in which case the ess sup is taken to be $-\infty$. Next, e.g.,

$$\text{ess} \limsup_{s \to t+} f(s)$$

is defined as the infimum of the ess sup of f on $(t, t+n^{-1})$ as $n \to \infty$; ess inf and ess lim inf are defined in a similar way. When ess lim $\sup_{s \to t+} f(s)$ and ess lim $\inf_{s \to t+} f(s)$ are equal we say that ess $\lim_{s \to t+} f(s)$ exists and is equal to the common value. We can of course define the latter directly but we need the other concepts below.

Some of the properties of ess $\lim_{s \to t+} f(s)$ are summarized in the next lemma, whose proof is omitted, being elementary analysis.

LEMMA 5.1. *Suppose that for every t in R, $\varphi(t) = \text{ess} \lim_{s \to t+} f(s)$ exists. Then φ is right continuous everywhere, $f = \varphi$ except for a set Z of measure zero, and we have*

$$\forall t: \quad \text{ess} \lim_{s \to t+} f(s) = \lim_{\substack{s \to t+ \\ s \notin Z}} f(s).$$

Finally, we have

$$\forall t: \varphi(t) = \lim_{\lambda \to \infty} \int_0^\infty \lambda e^{-\lambda s} f(t+s)\, ds = \lim_{\delta \downarrow 0} \frac{1}{\delta} \int_t^{t+\delta} f(s)\, ds. \tag{5.1}$$

The next two propositions resemble the main lemmas for separability of a stochastic process due to Doob [6].

LEMMA 5.2. *Suppose that $H \in \mathcal{B} \times \mathcal{F}$ (the product Borel field of $R \times \Omega$) and put for each $t \in R$:*

$$H(t) = \{\omega : (t, \omega) \in H\}.$$

Let

$$\Delta = \left\{\omega : \int_0^\infty 1_H(t, \omega)\, dt > 0\right\}$$

and let Z be an arbitrary subset of R with $m(Z) = 0$. Then there exists a countable dense subset $D = \{t_n, n \geq 1\}$ of R such that $D \cap Z = \emptyset$ and

$$P\{\Delta \triangle \bigcup_n H(t_n)\} = 0, \tag{5.2}$$

where "$\triangle$" denotes the symmetric difference.

Proof. By Fubini's theorem, $[R \times (\Omega - \Delta)] \cap H$ has $m \times P$ measure zero and there exists $Z' \subset R$ with $m(Z') = 0$ such that if $t \notin Z'$ then $P(H(t) \setminus \Delta) = 0$. Let

$$T = \{t \in R: P(H(t)) > 0\}$$

and consider the class of sets of the form

$$\bigcup_{t \in C} H(t),$$

where C is a countable dense subset of R, disjoint from $Z \cup Z'$. A familiar argument shows that there is a set in the class whose probability is maximal. Call this set Λ and we will show that $P(\Delta \setminus \Lambda) = 0$. Otherwise let $\Delta \setminus \Lambda = \Lambda_0$, $P(\Lambda_0) > 0$. Then $H \cap (R \times \Lambda_0)$ has strictly positive $m \times P$ measure by definition of Δ and Fubini's theorem. Hence by the same theorem there exists some $t \notin Z \cup Z'$ such that $P(H(t) \cap \Lambda_0) > 0$, which contradicts the maximality of Λ since $\Lambda \cup H(t)$ would be in the class above and have a strictly greater probability than Λ. Finally, by the definition of Z' and the choice of C, it is clear that $P(\Lambda \setminus \Delta) = 0$.

THEOREM 5.1. *Let $\{Y_t, t \in R\}$ be an extended real-valued Borel measurable stochastic process in $(\Omega, \mathcal{F}, P)$. There exists Ω_0 in $\mathcal{F}$ with $P(\Omega_0) = 1$ and a countable dense set D of R with the following property. For each $\omega \notin \Omega_0$ and every nonempty open interval I of R, we have*

(i) $\operatorname*{ess\,sup}_{t \in I} Y(t, \omega) = \sup_{t \in I \cap D} Y(t, \omega)$

(ii) $\operatorname*{ess\,inf}_{t \in I} Y(t, \omega) = \inf_{t \in I \cap D} Y(t, \omega)$.

Such a set D will be referred to as an "essential limit set for Y".

Proof. For each I with rational endpoints, consider the set

$$\{(t, \omega): t \in I; Y(t, \omega) < \operatorname*{ess\,inf}_{s \in I} Y(s, \omega) \text{ or } Y(t, \omega) > \operatorname*{ess\,sup}_{s \in I} Y(s, \omega)\}.$$

This has $m \times P$ measure zero by Fubini's theorem, hence there is a subset $Z(I)$ of I with measure zero such that if $t \in I - Z(I)$ then for almost every ω:

$$\operatorname*{ess\,inf}_{s \in I} Y(s, \omega) \leqslant Y(t, \omega) \leqslant \operatorname*{ess\,sup}_{s \in I} Y(s, \omega). \tag{5.3}$$

Let Z be the union of $Z(I)$ over all such I. Next, for each rational r, consider

$$H = \{(t, \omega): t \in I; Y(t, \omega) > r\}$$

and define Δ corresponding to H as in Lemma 5.2. It follows that there exists a countable dense set $\{t_n, n \geqslant 1\}$, disjoint form Z, such that (5.2) is true. Observe that Δ is the set of ω where $\operatorname{ess\,sup}_{t \in I} Y(t, \omega) > r$, while $\bigcup_n H(t_n)$ is the set of ω where $\sup_n Y(t_n, \omega) > r$.

Hence if we denote by D_1 the countable set obtained by uniting the sequences $\{t_n\}$ over all I and r, D_1 is disjoint from Z and we have for almost every ω:

$$\operatorname*{ess\,sup}_{t\in I} Y(t,\omega) \leqslant \sup_{t\in D_1\cap I} Y(t,\omega).$$

Similarly, there is a countable set D_2 disjoint from Z such that for almost every ω:

$$\operatorname*{ess\,inf}_{t\in I} Y(t,\omega) \geqslant \inf_{t\in D_2\cap I} Y(t,\omega).$$

Then if $D = D_1 \cup D_2$, we have for almost every ω and every I:

$$\inf_{t\in D\cap I} Y(t,\omega) \leqslant \operatorname*{ess\,inf}_{t\in I} Y(t,\omega) \leqslant \operatorname*{ess\,sup}_{t\in I} Y(t,\omega) \leqslant \sup_{t\in D\cap I} Y(t,\omega).$$

But since $D \cap Z = \varnothing$, the first and last inequalities above can be reversed by (5.3), proving the theorem.

It will appear in our later applications of Theorem 5.1 to Theorems 6.1 and 6.3 that we shall not need its full strength but merely the existence of a countable dense set D such that if almost all paths have left and right limits along D then they have essential left and right limits. Thus it is sufficient to have the equations in (i) and (ii) above replaced by "$\leqslant$" and "$\geqslant$" respectively. Doob has pointed out that Theorem 5.1 can be circumvented by arguing with separable versions, see the end of proof of Theorem 6.1.

6. The moderately strong Markov property of the reverse process

In this section we assume that the given process X is strongly Markovian relative to right continuous fields $\{\mathfrak{F}_t, t \geqslant 0\}$, whose paths are not only right continuous on $0 \leqslant t < \infty$ but also have left limits everywhere on $0 < t < \alpha$. Thus for each optional T, $t > 0$ and bounded $\mathcal{E}$-measurable f, we have almost surely

$$E\{f(X_{T+t}) \mid \mathfrak{F}_{T+}\} = P_t(X_T, f).$$

We shall use the "shift operator" θ in the usual way but we remind the reader that we are dealing with a process with a fixed initial distribution and not a family of processes starting at each x.

Let $\{Y_t, t \geqslant 0\}$ be the extended process with lifetime $\alpha^* = \alpha + \beta$ as defined in § 4. Let $\{\tilde{\mathfrak{F}}_t, t > 0\}$ be the Borel field generated by the reverse process $\{\tilde{Y}_t, t > 0\}$ and $\hat{P}_t$ its transition function. As we have seen, $\hat{P}_t$ acts like a transition function of $\tilde{X}$ as well. A random variable (or simply "time") T will be called "reverse-optional" iff for every $t > 0$, we have $\{T < t\} \in \tilde{\mathfrak{F}}_t$; it is "strictly" so iff $\{T < t\}$ is replaced by $\{T \leqslant t\}$. This distinction is necessary as the

fields $\tilde{\mathfrak{F}}_t$, unlike $\mathfrak{F}_t$, are not necessarily right continuous. The Borel fields $\tilde{\mathfrak{F}}_{T+}$ and $\tilde{\mathfrak{F}}_T$ are defined in the usual way as in § 1. T is said to be "reverse-predictable" iff there exists a sequence of reverse-optional times $\{T_n\}$ such that $T_n < T$ and $T_n \uparrow T$ almost surely; in this case we have

$$\tilde{\mathfrak{F}}_{T-} = \bigvee_n \tilde{\mathfrak{F}}_{T_n+},$$

where for an increasing family of Borel fields $\{\mathcal{G}_t, t>0\}$ and an arbitrary random variable T, $\mathcal{G}_{T-}$ is the Borel field generated by the class of sets of the form $\{T>t\} \cap \Lambda$ with $\Lambda \in \mathcal{G}_t$. See [13] for a general discussion of the notion.

We begin with a useful lemma, whose proof is omitted as being intuitively obvious and technically familiar.

LEMMA 6.1. *Let D be a countable dense subset of $R=[0, \infty)$. Let T be an optional time (with respect to $\{\mathfrak{F}_t\}$) with the following property. If $T(\omega) < \infty$ then there is a subset $C(\omega)$ of D such that for each t, the set $\{\omega: t \in C(\omega)\} \in \mathfrak{F}_t$ and for each $\delta > 0$,*

$$(T(\omega),\ T(\omega)+\delta) \cap C(\omega) \neq \varnothing. \tag{6.1}$$

Then there exists a sequence of strictly optional times $\{T_n\}$ such that for each n:

$$T_n(\omega) \in C(\omega), \quad T_n(\omega) > T(\omega)$$

and $T_n \downarrow T$ on $\{T<\infty\}$.

Let T be predictable and (6.1) *be replaced by*

$$(T(\omega)-\delta,\ T(\omega)) \cap C(\omega) \neq \varnothing, \quad \text{for } 0<T(\omega)<\infty.$$

Then a similar conclusion is true if "$>$" and "$\downarrow$" are replaced by "$<$" and "$\uparrow$", and $\{T<\infty\}$ by $\{0<T<\infty\}$.

THEOREM 6.1. *Let D be a countable dense subset of R, $t>0$ and $f \in C_K$. Then almost surely the path*

$$s \to \hat{P}_t f \circ \tilde{Y}_s,$$

has left and right limits along D everywhere. In particular, it has left and right essential limits everywhere.

Proof. Let $\{T_k\}$ be a sequence of D-valued strictly reverse-optional times decreasing to a limit T. Notice that on $\{T<\beta\}$, $\tilde{Y}(T_k) \in \{\Delta_1 \cup \Delta_2 \cup \Delta_3\}$ for all large enough k. Since $\tilde{Y}$ is Markovian as proved in § 4, the strong Markov property holds at any discrete strictly reverse-optional time such as T_k, hence

$$E\{f \circ \tilde{Y}(T_k+t);\, T_k > \beta \mid \tilde{\mathfrak{F}}_{T_k}\} = 1_{\{T_k>\beta\}}\, \tilde{P}_t f \circ \tilde{Y}(T_k). \tag{6.2}$$

Since the paths of X have left limits except possibly at α, those of $\tilde{Y}$ have right limits except possibly at β. Thus $f \circ \tilde{Y}(T_k+t)$ converges as $k \to \infty$ on the set $\{T \geqslant \beta\}$, while $\tilde{\mathfrak{F}}_{T_k} \downarrow \wedge_k \tilde{\mathfrak{F}}_{T_k}$. Therefore by Hunt's lemma cited in § 1, the first member of (6.2) converges almost surely. Similarly if $T_k \uparrow$. Writing for short

$$g = \tilde{P}_t f,$$

we have proved that almost surely

$$\lim g \circ \tilde{Y}(T_k) \tag{6.3}$$

exists for any monotone sequence $\{T_k\}$ as specified above. Note that if $T_k \uparrow \infty$, then $\tilde{Y}(T_k) = \tilde{\Delta}$ for all sufficiently large k.

Let $a < b$ and put

$$T' = \inf\{r \in D \colon g(\tilde{Y}_r) < a\},$$

$$T'' = \inf\{r \in D \colon g(\tilde{Y}_r) > b\},$$

where we may suppose that $0 \notin D$. Define inductively

$$S_0 = 0, \quad S_1 = T', \quad S_2 = S_1 + T'' \circ \theta_{S_1},$$

$$S_{2n-1} = S_{2n-2} + T' \circ \theta_{S_{2n-2}}, \quad S_{2n} = S_{2n-1} + T'' \circ \theta_{S_{2n-1}}, \quad n \geqslant 2.$$

These are all reverse-optional times not necessarily D-valued. It is possible that $S_0 = S_2$, but we have $S_n < S_{n+2}$ almost surely for $n \geqslant 1$. For otherwise on the set $\{S_n = S_{n+1} = S_{n+2}\}$ we have

$$\varliminf_{\substack{r \to S_n+ \\ r \in D}} g(\tilde{Y}_r) \leqslant a < b \leqslant \varlimsup_{\substack{r \to S_n+ \\ r \in D}} g(\tilde{Y}_r).$$

By Lemma 6.1 we can then construct D-valued, reverse-optional $\{T_k\}$ such that $T_k \downarrow S_n$ on the set above and

$$g \circ \tilde{Y}(T_k) \leqslant a, \quad g \circ \tilde{Y}(T_{k+1}) \geqslant b, \tag{6.4}$$

contradicting (6.3).

Next, we show that $S_n \to \infty$ almost surely. For on the set $\{S_n \uparrow S < \infty\}$ we can construct as before D-valued, strictly reverse-optional times $\{T_k\}$ such that $T_k \uparrow S$ and (6.4) holds, again contradicting (6.3). The fact that $S_n \uparrow \infty$ almost surely shows that there is no point in R at which the oscillation on the left or on the right of $g \circ \tilde{Y}_s$, $s \in D$, exceeds $b-a$. If we consider all rational pairs $a<b$, we conclude that $s \to g \circ \tilde{Y}_s$ must have left and right limits along D, everywhere in R. Taking D to be the essential limit set for $\tilde{Y}$ in Theorem 5.1,

we see that the existence of such limits is equivalent to the existence of left and right essential limits, q.e.d.

Instead of using Theorem 5.1 as in the last sentence above, we may conclude in the following way as suggested by Doob. Let Y' be a separable version of $\tilde{Y}$ with separability set D' such that almost surely $\tilde{Y}(s) = Y'(s)$ for all $s \in D$. Then almost all paths of Y' have left and right limits along D' and consequently by separability have left and right limits without restriction. By Fubini's theorem, almost every path $s \to \tilde{Y}_s(\omega)$ differs from the corresponding path $s \to Y'_s(\omega)$ on a set of s of Lebesgue measure zero. It follows that the former has essential left and right limits.

COROLLARY. *The assertion of the theorem is also true for almost every path*

$$s \to \hat{P}_t f \circ Y_s.$$

For the essential right limit, e.g., at s, of $\hat{P}_t f \circ Y_{\cdot}(\omega)$ is just the essential left limit of $\hat{P}_t f \circ \tilde{Y}_{\cdot}(\omega)$ at $\alpha^*(\omega) - s$, since $\tilde{Y}_s = Y_{\alpha^* - s}$ for all $0 < s \leqslant \alpha^*$.

Recalling that (P_t) is the transition function of X, we put

$$R_\lambda = \int_0^\infty e^{-\lambda t} P_t \, dt, \quad \lambda > 0$$

as its resolvent. We shall use this operator only as a familiar way of integral averaging. A set in $\mathcal{E}$ which is hit by X with probability zero will be called "polar".

THEOREM 6.2. *Let g be bounded, $\mathcal{E}$-measurable and suppose that almost surely the path*

$$s \to g(X_s)$$

has essential right limits everywhere. Then the following limit

$$\hat{g}(x) \overset{\text{def}}{=} \lim_{\lambda \to \infty} \lambda R_\lambda g(x) \tag{6.5}$$

exists except possibly for a polar set; and we have almost surely

$$\forall s \geqslant 0: \operatorname*{ess\,lim}_{r \to s+} g(X_r) = \hat{g}(X_s). \tag{6.6}$$

Proof. Put

$$Z_s = \operatorname*{ess\,lim}_{r \to s+} g(X_r);$$

without loss of generality we may suppose that this limit exists everywhere on Ω. The process $(s, \omega) \to Z(s, \omega)$ is measurable since by Theorem 5.1 the essential limit may be replaced by that on a countable set. It is also right continuous and hence well-measurable

in the sense of Meyer. Let us complete the definition of $\hat{g}$ by setting it to be a constant greater than an upper bound of g wherever the limit in (6.5) fails to exist. Since $\hat{g} \in \mathcal{E}$, $\hat{g}(X)$ is well-measurable with X, and consequently the set

$$H = \{(s, \omega): Z(s, \omega) \neq \hat{g}(X(s, \omega))\} \tag{6.7}$$

is well-measurable. Let $\Pi(H)$ be its projection on Ω. If $P(\Pi(H)) > 0$, then a theorem by Meyer [12; p. 204] aserts that there exists an optional time T such that $P\{T < \infty\} > 0$ and $(T(\omega), \omega) \in H$ so that $Z_T \neq \hat{g}(X_T)$ on $\{T < \infty\}$. But (almost surely in the third and fifth equations below)

$$\begin{aligned} Z_T &= \operatorname*{ess\,lim}_{s \to T+} g(X_s) = \lim_{\lambda \to \infty} \int_0^\infty \lambda e^{-\lambda t} g(X_{T+t})\, dt \\ &= \lim_{\lambda \to \infty} E\left\{\int_0^\infty \lambda e^{-\lambda t} g(X_{T+t})\, dt \,|\, \mathfrak{F}_{T+}\right\} = \lim_{\lambda \to \infty} \int_0^\infty \lambda e^{-\lambda t} E[g(X_{T+t}) \,|\, \mathfrak{F}_{T+}]\, dt \\ &= \lim_{\lambda \to \infty} \int_0^\infty \lambda e^{-\lambda t} P_t g(X_T)\, dt = \lim_{\lambda \to \infty} \lambda R_\lambda g(X_T) = \hat{g}(X_T), \end{aligned}$$

which is a contradiction. Hence $P\{\Pi(H)\} = 0$ and (6.6) follows. Let A denote the set of x for which the limit in (6.5) fails to exist. Then on $\{T_A < \infty\}$ there exists $s \geq 0$ such that $Z_s < \hat{g}(X_s)$. Thus $\{T_A < \infty\} \subset \Pi(H)$ and A is a polar set.

Recalling that X is an initial portion of Y, we may apply Theorem 6.2 to $g = \hat{P}_t f$ on account of the Corollary to Theorem 6.1. Thus for each $t > 0$ there exists a polar set A such that for all $f \in C_K$, $x \in E - A$, the following limit

$$\hat{P}_t f(v) \overset{\text{def}}{=} \lim_{\lambda \to \infty} \lambda R_\lambda (\hat{P}_t f)(x) \tag{6.8}$$

exist. Set $\hat{P}_t f(x) = 0$ if $x \in A$. The operator $\hat{P}_t$ may be extended to a kernel in the usual way. We state this as follows.

COROLLARY. *For each $t > 0$ and $f \in C_K$, we have almost surely*

$$\forall s \geq 0: \operatorname*{ess\,lim}_{r \to s+} \hat{P}_t f \circ X_r = \hat{P}_t f \circ X_s. \tag{6.9}$$

In particular, $s \to \hat{P}_t f \circ X_s$ is almost surely right continuous.

The last sentence above would amount to the fine continuity of $x \to \hat{P}_t f(x)$ in the customary set-up where the process X is allowed to start at an arbitrary x.

THEOREM 6.3. *Let T be a reverse-predictable time. Then each $t > 0$ and $f \in C_K$, we have*

$$E\{f \circ \tilde{Y}(T + t) \,|\, \tilde{\mathfrak{F}}_{T-}\} = \hat{P}_t f \circ \tilde{Y}(T). \tag{6.10}$$

Proof. Since $\tilde{Y}_s = Y_{\alpha^*-s}$, it follows from (6.9) extended trivially to Y that we have almost surely

$$\forall s > 0: \operatorname*{ess\,lim}_{r \to s-} \hat{P}_t f \circ \tilde{Y}_r = \hat{P}_t f \circ \tilde{Y}_s. \tag{6.11}$$

Since T is reverse-predictable, there exists reverse-optional times $\{T_n\}$ such that $T_n < T$, $T_n \uparrow T$ almost surely. It follows from Hunt's lemma and the left continuity of the paths of $\tilde{Y}$ that

$$E\{f \circ \tilde{Y}(T_n + t) \mid \tilde{\mathfrak{F}}_{T_n+}\} \to E\{f \circ \tilde{Y}(T+t) \mid \tilde{\mathfrak{F}}_{T-}\}. \tag{6.12}$$

Let D be an essential limit set for the process $\{P_t f \circ \tilde{Y}_s,\ s \geq 0\}$. By Lemma 6.1 we can find an increasing sequence of reverse-optional times $\{T'_n\}$, D-valued and such that $T_n \leq T'_n < T$ for all n. The strong Markov property holds at T'_n since $\tilde{Y}$ is Markovian, so that the left member of (6.12), after T_n is replaced by T'_n, becomes $\hat{P}_t f \circ \tilde{Y}(T'_n)$. By (6.11), the latter converges as $n \to \infty$ to

$$\operatorname*{ess\,lim}_{s \to T-} \hat{P}_t f \circ \tilde{Y}(s) = \hat{P}_t f \circ \tilde{Y}(T).$$

Thus (6.12) becomes (6.10), q.e.d.

THEOREM 6.4. *The equation* (6.10) *remains true if $\tilde{Y}$ is replaced by $\tilde{X}$ and T is predictable with respect to $\{\mathcal{G}_t, t > 0\}$ where $\mathcal{G}_t$ is the Borel field generated by $\{\tilde{X}_s, 0 < s \leq t\}$. In particular $\tilde{X}$ is a homogeneous Markov process with $\{\hat{P}_t, t > 0\}$ as transition function in the loose sense.*

Proof. Recall the β in § 4 such that $\tilde{Y}_{\beta+t} = \tilde{X}_t$, $t > 0$. If T is predictable relative to $\mathcal{G}_t$ as stated, then $\beta + T$ is reverse-predictable. Furthermore, we have

$$\mathcal{G}_{T-} \subset \tilde{\mathfrak{F}}_{(\beta+T)-}. \tag{6.13}$$

To see this we observe first that $\tilde{Y}_{\beta+t} \in \tilde{\mathfrak{F}}_{\beta+t}$ be left continuity of paths, hence $\mathcal{G}_t \subset \tilde{\mathfrak{F}}_{\beta+t}$ and so if $\Lambda \in \mathcal{G}_t$, then for each q, $\{q > \beta + t\} \cap \Lambda \in \tilde{\mathfrak{F}}_q$. Hence for each t,

$$\{T > t\} \cap \Lambda = \bigcup_{q \in Q} [\{\beta + T > q\} \cap \{q > \beta + t\} \cap \Lambda]$$

belongs to $\tilde{\mathfrak{F}}_{(\beta+T)-}$ since each member of the union does, by definition of the field. This proves (6.13) by definition of $\mathcal{G}_{T-}$. Substituting $\beta + T$ for T in (6.10) we obtain

$$E\{f \circ \tilde{X}(T+t) \mid \tilde{\mathfrak{F}}_{(\beta+T)-}\} = \hat{P}_t f \circ \tilde{X}(T);$$

together with (6.13) this implies the first assertion of the theorem. Now take T to be a constant $t_0 > 0$, and observe that as $\mathcal{G}_{t_0-} = \mathcal{G}_{t_0}$ by the left continuity of paths, the resulting equation then implies the second assertion of the theorem.

References

[1]. BLACKWELL, D., On a class of probability spaces. *Proceedings of the Third Berkeley Symposium on Math. Stat. and Prob.*, Vol. 2, pp. 1–6. University of California Press, 1956.

[2]. BLUMENTHAL, R. M. & GETOOR, R. K., *Markov Processes and Potential Theory*. Academic Press, 1968.

[3]. CARTIER, P., MEYER, P. A. & WEIL, M. Le retournement du temps: compléments à l'exposé de M. Weil. *Séminaire de Probabilités II, Université de Strasbourg*, pp. 22–33. Springer-Verlag, 1967.

[4]. CHUNG, K. L., On the Martin boundary for Markov chains. *Proc. Nat. Acad. Sci.*, 48 (1962) 963–968.

[5]. —— *Markov Chains with Stationary Transition Probabilities.* 2nd ed. Springer-Verlag, 1967.

[6]. DOOB, J. L., *Stochastic Processes.* Wiley & Sons, 1953.

[7]. HUNT, G. A., Markoff processes and potentials III. *Ill. J. Math.*, 2 (1958), 151–213.

[8]. —— Markoff chains and Martin boundaries. *Ill. J. Math.*, 4 (1960), 313–340.

[9]. —— *Martingales et Processus de Markov.* Dunod, 1966.

[10]. IKEDA, N., NAGASAWA, M. & SATO, K., A time reversion of Markov processes and killing. *Kodai Math. Sem. Rep.*, 16 (1964), 88–97.

[11]. KUNITA, H. & WATANABE, T., On certain reversed processes and their applications to potential theory and boundary theory. *J. Math. Mech.*, 15 (1966), 393–434.

[12]. MEYER, P. A., *Probabilités et potentiel.* Hermann, 1966.

[13]. —— Guide détaillé de la théorie "générale" des processus. *Séminarie de Probabilités II, Université de Strasbourg*, pp. 140–165. Springer-Verlag, 1968.

[14]. —— Processus de Markov: la frontière de Martin. *Séminaire de Probabilités III, Université de Strasbourg.* Springer-Verlag, 1968.

[15]. NAGASAWA, M., Time reversions of Markov processes. *Nagoya Math. J.*, 24 (1964), 177–204.

[16]. NELSON, E., The adjoint Markov process. *Duke Math. J.*, 25 (1958), 671–690.

[17]. WEIL, M., Retournement du temps dans les processus Markoviens. Résolventes en dualité. *Séminaire de Probabilités I, Université de Strasbourg*, pp. 166–189. Springer-Verlag, 1967.

Received May 13, 1969.

Actes, Congrès intern. Math., 1970. Tome 2, p. 499 à 505.

D 5 - THÉORIE DU POTENTIEL
PROCESSUS DE MARKOV

BOUNDARY BEHAVIOR OF MARKOV CHAINS AND ITS CONTRIBUTIONS TO GENERAL PROCESSES

by Kai Lai CHUNG (1)

In contemporary studies of homogeneous Markov processes on a topological space, under the name of Hunt or standard process, it is assumed that the only discontinuities of (almost all) sample functions are jumps, for all time or up to the lifetime of the process, respectively. If the same assumption is made on a Markov chain, where the state space is discrete and may be labeled by the integers, this results in a rather trivial situation long since "solved". If other types of discontinuity are allowed, then the typical sample function will have infinity as a limiting value when such a discontinuity is approached, from one or both directions of time. Various ways of reaching and returning from infinity should then be distinguished, and the consequent ramification has been called a boundary in analogy with classical potential theory. The problem is then to set up a suitable boundary and investigate the behavior of the sample functions relative to it. For the proper object of study of stochastic processes is the collection of sample functions or paths – it is through the underpinning a groundwork of paths that modern probability theory makes its most original contribution to mathematics(2).

It is easy to give a formulation in a more general context. For instance, given a standard process X with its lifetime T, we may inquire after the structure of all homogeneous Markov processes X^* with X as its initial portion and hence (if some form of strong Markov property is to hold for X^*) as a germinal constituent, in the sense that the paths should behave as they do in X away from a certain boundary, or again that X^* should be decomposable into X and a boundary derived from X. In the case of Markov chains, X may be a minimal chain (see [3] for this and other standard terminology) whose paths are of the trivial type mentioned above, controlled by an initial derivative matrix Q which will be assumed to be conservative. This leads to the so-called complete construction problem : given Q, to find all transition semigroups P such that $P'(0) = Q$. This formulation is popular among those mathematicians who wish to stake out an easily stated analytic pro-

---

(1) Research supported by the Office of Scientific Research Office of the United States Air Force, under AFOSR contract F44620-67-C-0049.

(2) In this vein it is curious to compare the works on stochastic processes by Lévy and Doob on one hand, and Feller on the other. Behind Feller's analytic doing, however, there always lurks his thinking in terms of paths.

blem devoid of probabilistic content. However, the way to all construction is of course an adequate understanding of the fundamental structure of the would-be-constructed object, as any school child who has figured out a regular hexagon should know. From my point of view, therefore, the main thing is to describe the evolution of time of a process, in other words to trace a typical path through its ups and downs at the boundary. Definitive results are known only in the case where the exit boundary is finite or discretely countable (see [2], [4], [9], [13]). It is probably inevitable that as more general cases are treated, the finer details will become blurred, and it is not clear what kind of meaningful results can be achieved in utter generality.

We have yet to define a boundary. The word brings to mind several cognate names in other contexts, and the tendency is strong nowadays to fit a ready-made blueprint onto an emerging situation. This has its obvious advantages, but one runs the risk of losing sight of a green field because of skyscrapers and superhighways, figuratively speaking. Since the specific case that can be handled is simple enough, I choose to describe it intuitively and without punctilio. Assuming then that the minimal chain is transient as we may, it is known (after Blackwell) that the path will ultimately enter and remain in an invariant set, namely a set A such that $\liminf\{X_t \in A\} = \limsup\{X_t \in A\}$ almost surely (a.s.) as t increases to T, where T is the lifetime of the minimal chain and is also the first boundary hitting time of the whole process. Note that T is a predictable time (see [10]) as the limit of a sequence of strictly increasing jump times. Now if we assume that there is only a finite number of atomic invariant sets that can be reached in finite time, we will identify each of them with a boundary point and say that $X(T-) = b$ if b corresponds to the set A above. The path has thus crossed the boundary **B** at b and the question is what it does thereafter. A classification of boundary points into "sticky" and "nonsticky" ones will be made. The boundary point b is sticky iff after first hitting b the path must a.s. hit it infinitely many times immediately afterward. b is nonsticky iff after first hitting b the path must a.s. not hit it again for a strictly positive time. This dichotomy is a form of special zero-one law (which does not hold in general as for a standard process). The distinction is important because in the sticky case it precludes the possibility of considering successive hits after the first. To circumvent this difficulty, a simple but crucial device is used as follows. Instead of successive hitting, we consider the successive "switching" of boundary encounter ; namely, after the first hit at T_1 we define T_2 to be the first time (an infimum in the usual manner) the path hits a boundary point different from that of the first hit, T_3 to be the first time the path hits a boundary point different from that of the second hit (but may of course be the same as the first hit), and so forth. This recursive definition is complicated by the possibility that a switch may occur instantly, for instance T_1 may equal T_2 if $X(T_1-) = b_1$ but the path hits some other boundary point b_2 in $(T_1, T_1 + \epsilon)$ for every $\epsilon > 0$. This can happen only if b_1 is nonsticky and b_2 is sticky, hence an instant switch cannot happen twice in succession. Thus the sequence of switching times T_n must strictly increase at least every other time. They cannot accumulate to a finite limit, for at such a time the path would have to oscillate between distinguishable boundary points, which is a.s. impossible by a martingale argument. Thus either $T_n = \infty$ from some random value of n on,

or $T_n \uparrow \infty$ a.s. We have therefore partitioned the time axis $[0, \infty)$ into disjoint abutting subintervals :

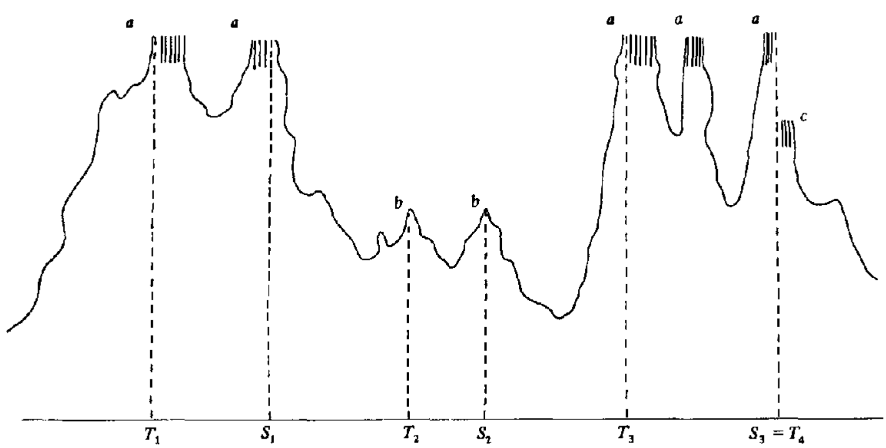
$$[0, T_1) \cup [T_1, T_2) \cup [T_2, T_3) \cup \ldots$$

in each of which at most one particular boundary point can be hit. Such a reduction to one is clearly desirable.

If $X(T_1 -) = a$ we call the process in $[T_1, \infty)$ the post-a process ; and we call the process in $[T_1, T_2)$ the sub-a process. It transpires by virtue of a strengthened Markov property applicable at the boundary, to be discussed below, that whenever $X(T_n-) = a$ the process in $[T_n, \infty)$ is stochastically equivalent to the post-a process, and the process in $[T_n, T_{n+1})$ is equivalent to the sub-a process. Clearly, a post-a process is just the process starting at the boundary point a, and a sub-a process is this process killed at $\mathbf{B} - \{a\}$. Now consider a typical nondegenerate subinterval and denote it by $[T, T')$. We know by definition that

$$X(T-) = a, \quad X(T'-) = b$$

where $a \neq b$; we know also that the path does not hit any boundary point except possibly a in the interval. Let the last hit (defined as a supermum) of a be S. This may coincide with T (which can happen only if a is nonsticky) or with T' (which can happen only if b is sticky). The following picture illustrates the various possibilities :



We shall indicate how the basic quantities for the process can be derived from the preceding description of the paths in three stages. The reader will have to consult [2] or [4] for formal definitions of the symbols below. From the first hitting of the boundary, we get

$$(1) \qquad p_{ij}(t) = f_{ij}(t) + \sum_{a \in \mathbf{B}} \int_0^t l_i^a(s)\, \xi_j^a(t-s)\, ds \ ;$$

where $\Pi = (p_{ij})$ is the transition function of the whole chain, $\Phi = (f_{ij})$ that of the minimal chain, l^a is the first hitting time density at a, which is an exit law

for Φ, ξ^a is the normalized entrance law for the post-a process. From the switching, we get

$$(2) \qquad \xi_j^a(t) = \rho_j^a(t) + \sum_{b\in B} \int_0^t F^{ab}(ds)\, \xi_j^b(t-s) \;;$$

where ρ^a is the entrance law of the sub-a process (which can be normalized by adjoining the usual death point), $F^{ab}(dt)$ is the switching time distribution from a to b. From the last exit in each subinterval, we get

$$(3) \qquad \rho_j^a(t) = \int_0^t E^a(ds)\, \eta_j^a(t-s) \;;$$

where $E^a(dt)$ is the distribution of S in the sub-a process (provided a is not a recurrent trap), and η^a is the canonical Φ-entrance law linked to the exit a, to be discussed below.

Putting together these three formulas and introducing Laplace transforms for conciseness, we obtain the complete decomposition formula

$$(4) \qquad \hat{p}_{ij}(\lambda) = \hat{f}_{ij}(\lambda) + \sum_{a\in B}\sum_{b\in B} \hat{l}_i^a(\lambda)\, [(1-\hat{F}(\lambda))^{-1}\hat{E}(\lambda)]_{(a,b)}\, \hat{\eta}_j^b(\lambda) \;;$$

where $[1-\hat{F}(\lambda)]^{-1}\hat{E}(\lambda)$ is a matrix on $\mathbf{B}\times\mathbf{B}$. This formula is the key to the construction problem mentioned earlier.

For a thorough analysis of the movement of the paths, certain critical combinations of the quantities above, and some new ones such as Π-exit laws, should be introduced. These become quite technical and so rather than going into them I will now discuss some of the ideas arising from this boundary study which will be found useful, indeed has already been, in the general theory of Markov processes.

The very first step in crossing the boundary involves a form of strengthened Markov property, specifically : if T is the first boundary hitting time, $\mathcal{F}_{T-}$ the strict pre-T field, $\mathcal{F}'_T$ the post-T field, then we have for every $\Lambda\in\mathcal{F}_{T-}$ and $M\in\mathcal{F}'_T$:

$$P\{\Lambda\cap M \mid X(T-)\} = P\{\Lambda \mid X(T-)\}\, P\{M \mid X(T-)\} \;;$$

or equivalently for every $t>0$ and A a measurable set of the state space :

$$P\{X(T+t)\in A \mid \mathcal{F}_{T-}\} = P_t(X(T-), A) \;;$$

where $P_t(x, A)$ is the usual notation for a transition probability. Observe that this differs from the usual strong Markov property in that $T-$ replaces $T+$ everywhere. We recall that such a property is known to hold for a homogeneous Markov chain in its right-lower-semi-continuous version, whenever $X(T+)$ belongs to the state space (see [1], [3]). This is not necessarily the case at a boundary hitting time, whence the need for a new departure(1). Although much work was done

---

(1) There are brief remarks about the boundary in [1], and some illustrations of the problem in [3]. In retrospect, the approaches to a boundary theory for Markov processes have progressed rapidly.

in the early days on the strong Markov property, this seems to be the first entry (see [2]) of a left-oriented version to deal with changed circumstances(1). In fact, although the right field $\mathscr{F}_{T+}$ has been in use since the beginning of Hunt's theory, its natural companion $\mathscr{F}_{T-}$ and the concomitant predictable time such as the T in question was a more recent addition (see [5] and [10]). Later it turned out that this is the form of Markov property, named "moderate Markov property", that holds in general when a strong Markov process with right continuous paths having left limits is reversed in time (see [6] and [11]). It is not a consequence of the other, right-oriented form even when both are meaningful, but it holds for a Hunt process as well, from which quasi-left continuity follows at once. There is now reason to think that the moderate Markov property, rather than the customary quasi-left continuity, is entitled to the status of a preliminary hypothesis. Details of this suggestion will be developed elsewhere. (2)

An interesting case is that of an instant switch already mentioned. On the set where the first hitting of a is also an incipient hitting of b, the usual strong Markov property also holds as if $X(T+) = b$. At least in some compactification (see [15]), nonsticky boundary points coincide with branch points and the instant switch becomes a jump from a to b. Now the existence of branch points is known from abstract considerations (see [14]), but the boundary theory furnishes genuine examples of them so that their admittance to the general theory seems imminent. Instant switch from a last exit time, rendering the possibility of $T = T'$ in the discussion above, is a related phenomenon, the difference being that such a time is inaccessible instead of predictable.

Under our hypotheses, the Φ-exit laws l^a and the Π-entrance laws ξ^a are immediately definable from their probabilistic meanings. An essential difficulty, analytically as well as stochastically, is to find the Φ-entrance laws η^a. In the approach sketched above, these are picked out, so to speak, by the paths themselves, one for each exit. (There is no need of an entrance boundary, even as to its existence, although this may be a good thing to have (see [8]).) This derivation depends on the important observation that the potential of the sub-a process is finite, namely :

$$\forall j : e_j^a = \int_0^\infty p_j^a(t)\, dt < \infty ;$$

except when a is a recurrent trap in which case the e^a below is to be replaced by a quasi-stationary measure for the post-a process (identical to the sub-a process). Now e^a is an excessive measure for Φ and the excess has a continuous derivative which is precisely η^a :

$$e^a - e^a\,\Phi(t) = \int_0^t \eta^a(s)\, ds .$$

As a hindsight, it can be shown that $\eta^a\ (t)$ is also the limit

$$\lim_{s \downarrow 0} \frac{\xi^a(s)\,\Phi(t-s)}{1 - \langle \xi^a(s)\,, L^a(\infty) \rangle} .$$

---

(1) Compactifiers were of course hellbent on regularization to an old pattern, and ignored the opening to the left, but this has now been noticed (see [7]).

(2) At the *Convegno sul Calcolo delle probabilità* in Rome, March 1971.

This is intuitively more suggestive but perhaps conceals a fundamental limiting process. The method of converting a generally infinite potential for the post-a process $\int_0^\infty \xi^a(t)\,dt$ into the generally finite sub-a potential may be worth investigation. It is done by imposing a taboo set (here the boundary set $B - \{a\}$), a familiar device in Markov chain theory. The standard method of considering λ-potentials is of course just the resolvent theory, which has proved to be such a powerful tool. But calculations with resolvents tend to be purely algebraic manipulations, since it corresponds to a killing at an exponential time totally independent of and therefore alien to the process. By contrast, killing under a suitable taboo leads to simple probabilistic interpretations and more easily recognizable results(1). I do not know the scope of applicability of such an alternative for the general theory, but submit that we be on the look-out for it.

The idea of a last exit time from the boundary plays a curious role in the final step of the decomposition, expressing the entrance law of a one-boundary process in terms of that of a no-boundary process. In Markov chain theory, the last exit time from an ordinary state (particularly when it is instantaneous) is known to bear tricky consequences such as the differentiability of the transition function. A last exit being a first entrance when the sense of time is reversed, it should not be surprising that it figures prominently in the behavior of the paths, and its deeper impact is presumably due to an implicit reversal. Thus the true meaning of (3) is through a reversed viewing of the sub-a process $[T, T')$ from the terminal end T', so that the last exit time S becomes the lifetime of the minimal chain of the reverse subprocess. Indeed a final dénouement is obtained when the reverse subprocess is reversed again to retrieve the original subprocess. This doubling-back practice is by no means wasteful, as it shows up some fine features which are obscured by one-way thinking. In particular, one sees that the process from S to T' is also Markovian (although as stressed by Meyer, it is not a "subprocess" as defined by Dynkin), as well as the process from T to S, and furthermore there is stochastic independence between these two portions relative to their common epoch S. This is the reason for (3)(2). In general, a last exit time is both a death time and a birth time, and the notion has now been generalized to "co-optional" and "co-terminal" in [12], in the same sense that historically a first entrance time generalizes to "optional" and "terminal". The ensuing symmetry or duality with a concatenation should prove fertile.

It is well known that Hunt constructed the Martin boundary for a discrete parameter Markov chain by considering a sequence of last exit times from finite subsets swelling up to the state space. For a continuous parameter minimal chain this can be done at one swoop by reversing from its (finite) lifetime (see [3]).

---

(1) This remark applies to reversing from a finite lifetime, see below.

(2) Other proofs of (3) have been given based on a local time $A(t)$ at the boundary point a (see [9], [16]) and going back to an analysis by Neveu. This is not surprising, since $E^a(t) = \mathrm{E}\{S \leqslant t\} = \mathrm{E}\{A(t)\}$ in our notation, where the last two E's are expectations for the sub-a process. But so far the intuitive meaning has not been made clearer by this method than by considering the last exit time and reversing the time.

The same ideas were used in [6] to reverse a general homogeneous Markov process to obtain a homogeneous Markov process. By insisting on following the paths faithfully and refraining from forcing them into any preconceived pattern, it is possible to set things on a natural course. We obtain thus a pair of homogeneous Markov processes *in reverse* sharing the same collection of sample functions with the arrow of time pointing in opposite directions. This entails two Markovian semi-groups in duality but enriched with the common structure of the paths. Much more needs be done to substantiate this "reversal" (vs. "dual") point of view, but I think it is a good instance where the general theory can learn from the chains.

REFERENCES

[1] CHUNG K.L. — Continuous parameter Markov chains, *Proc. International Congress of Mathematicians*, 1958, p. 510-517.

[2] CHUNG K.L. — On the boundary theory for Markov chains, *Acta. Math.*, 110. 1963, p. 19-77; 115, 1966, p. 111-163.

[3] CHUNG K.L. — *Markov chains with stationary transition probabilities*, Second Edition, Springer-Verlag, 1967.

[4] CHUNG K.L. — *Lectures on boundary theory for Markov chains, Princeton University* Press, 1970.

[5] CHUNG K.L. and DOOB J.L. — Fields, optionality and measurability, *Amer. J. Math.*, 1965, p. 397-424.

[6] CHUNG K.L. and WALSH J.B. — To reverse a Markov Process, *Acta Math.*, 123, 1969, p. 225-251.

[7] DOOB J.L. — Compactification of the discrete state space of a Markov process, *Z. Wahrscheinlichkeitstheorie verw. Geb.*, 10, 1968, p. 236-251.

[8] DYNKIN E.B. — General boundary conditions for denumerable Markov processes, *Teor. Veroyatnost. i Primenen*, 12, 1967, p. 222-257.

[9] LAMB Charles W. — Boundary theory for Markov chains, *Doctorate Dissertation*, University of Illinois, 1969.

[10] MEYER P.A. — Guide détaillé de la théorie « générale » des processus, *Séminaire de probabilité*, II, p. 140-165, Springer-Verlag, 1968.

[11] MEYER P.A. — Le retournement du temps d'après Chung et Walsh, *Séminaire de probabilités*, V (to appear).

[12] MEYER P.A., SMYTHE R.T. and WALSH J.B. — *Birth and death of Markov processes*, Sixth Berkeley Symposium (to appear).

[13] PITTENGER A. — Boundary decomposition of Markov processes, *Doctorate Dissertation*, Stanford University, 1967.

[14] RAY D. — Resolvants, transition functions and strongly Markovian processes, *Annals of Math.*, 70, 1959, p. 43-72.

[15] WALSH J.B. — The Martin boundary and completion of Markov chains, *Z. Wahrscheinlichkeitstheorie verw. Geb.*, 14, 1970, p. 169-188.

[16] WILLIAMS D. — Fictitious states, coupled laws and local time, *Z. Wahrscheinlichkeitstheorie verw. Geb.*, 11, 1969, p. 288-310.

Stanford University
Dept. of Mathematics,
Stanford
California 94305 (USA)

Reprinted from *Amer. Math. Monthly* **79** (1972), pp. 867–877.

CRUDELY STATIONARY COUNTING PROCESSES

KAI LAI CHUNG, Stanford University

1. Introduction. The theorems by Khintchine, Korolyuk, and Dobrushin in the theory of stationary point processes are basic and simple theorems. Korolyuk's theorem was originally derived from the Palm-Khintchine formulas; a direct proof was given in Cramér-Leadbetter [1]. Its real simplicity seems to be obscured by the slightly complicated presentation of the proof. The same may be said of the proof of Dobrushin's theorem involving an unnecessary contraposition as well as some epsilonics. Both results become quite transparent when dealt with by standard methods of measure and integration in sample space. After all, these are problems of probability theory and nowadays students spend a lot of time learning this kind of "abstract" set-up. It would be a pity not to use the knowledge so acquired in straightforward situations such as these theorems. In doing so we arrive at certain natural extensions which seem to put the results in proper perspective. The results in R^d, obtained by the same method, seem to be new.

The reader is referred to Leadbetter [3] for another simple approach, which came belatedly to our attention.

2. Definitions and statements. Let $(\Omega, \mathscr{F}, P)$ be a probability space. Each ω in Ω is a set $S(\omega)$ of points on $R = (-\infty, +\infty)$ endowed with "multiplicity," namely a positive integer attached to the point. A point with **multiplicity** m is counted as m ordinary points at distance zero to each other; it will be called a **multiple point** when $m \geqq 2$. The fundamental assumptions are as follows:

(A) For each finite interval I in R, the "number" of points in $S(\omega) \cap I$, counted with their multiplicities, is finite. This number will be denoted by $N(I, \omega)$; as usual $N(I)$ is the function $\omega \to N(I, \omega)$.

(B) The function

$$(s, t, \omega) \to N([s, t]; \omega),$$

where $s \leqq t$ is measurable with respect to the product field $\mathscr{B} \times \mathscr{B} \times \mathscr{F}$ where $\mathscr{B}$ is the Euclidean Borel field on R.

It follows that for each interval I, $N(I)$ is a random variable. We do not define $N(\cdot)$ for other sets than intervals.

The collection $\{N(I, \omega)\}$ with I ranging over intervals and ω over Ω, will be called a **counting process** on R. It is said to be **crudely stationary** iff whenever I and J

Since receiving his Princeton Doctor's Degree, Kai Lai Chung has held positions at Princeton, Cornell, Columbia, Syracuse, Chicago, University of Illinois, and Stanford. He has spent leaves abroad at the University of Strasbourg and the ETH Zurich. His main research interest is probability and he is the author of *Markov Chains*, *Course in Probability Theory*, and *Boundary Theory*. *Editor.*

are two compact intervals of equal length the random variables $N(I)$ and $N(J)$ have the same distribution. The same will then be true for any two finite intervals of equal length, whether closed or open or half-open and possibly degenerate, by Proposition 1 below. Indeed the sole purpose of the formulation above is to bring that proposition into question. The adjective "crude" is used to distinguish it from "strict," which requires much more (see [1], 3.8).

An equivalent formulation is to define an integer-valued stochastic process $\{X(t,\omega);\, t \in R,\, \omega \in \Omega\}$ as follows:

$$X(t,\omega) = \begin{cases} N([0,t),\omega) & \text{if } t \geqq 0, \\ -N([t,0),\omega) & \text{if } t < 0. \end{cases}$$

For each ω, $t \to X(t,\omega)$ is then a right continuous purely jumping non-decreasing function. The set of its jump-points is $S(\omega)$ and the size of jump at each point is its multiplicity. If X has strictly stationary increments in the usual sense, then the **increment process** $\{N(I,\omega)\}$ will be not only crudely, but even strictly stationary. While the conversion to X has the advantage of making a counting process into a more standard object, the language and notation for N is slightly more direct and so preferred here. We begin by settling a small point, which in the strictly stationary case follows from the fact that the Borel-Lebesgue measure is the unique translation-invariant measure on R, apart from a constant factor.

We use E below to denote the mathematical expectation, and write, e.g., $\{N([t,t]) = 0\}$ for $\{\omega \mid N([t,t],\omega) = 0\}$.

PROPOSITION 1. *For each degenerate interval* $[t,t]$ *we have*

$$P\{N([t,t]) = 0\} = 1. \tag{1}$$

Proof. The set

$$H = \{(t,\omega) \mid N([t,t],\omega) \neq 0\} = \{(t,\omega) \mid t \in S(\omega)\}$$

belongs to $\mathscr{B} \times \mathscr{F}$ by (B). Integrating its indicator 1_H over $[0,1] \times \Omega$ and applying Fubini's theorem, we obtain in view of (A):

$$0 = \int_\Omega 0\, P(d\omega) = \int_\Omega \left[\int_{[0,1]} 1_H(t,\omega)\, dt\right] P(d\omega)$$

$$= \int_{[0,1]} \left[\int_\Omega 1_H(t,\omega)\, P(d\omega)\right] dt = \int_{[0,1]} E\{N([t,t])\}\, dt.$$

By crude stationarity, $t \to E\{N([t,t])\}$ is a constant c, where $0 \leqq c \leqq +\infty$; hence $c = 0$ which is equivalent to (1).

PROPOSITION 2. *Either* (i) $E\{N(I)\} = \infty$ *for every non-degenerate* I; *or* (ii)

$E\{N(I)\} < \infty$ for every finite I. In case (ii), *for any sequence $\{I_n\}$ such that $I_n \downarrow$ and $|I_n| \to 0$ (where $|I|$ = length of I), we have*

$$P\{\lim_n N(I_n) = 0\} = 1. \tag{2}$$

Furthermore, in either case we have for every $t \geqq 0$:

$$E\{N([0,t])\} = E\{N([0,1])\}\, t \tag{3}$$

provided we set $\infty \cdot 0 = 0$.

Proof. Observe that (2) is false in case (i) even when $I_n \downarrow \varnothing$. The rest follows from crude stationarity, dominated convergence and Proposition 1, and we omit the details of a familiar argument.

From now on we shall write

$$N(t) = N([0,t]),\ \mu(t) = E\{N(t)\},\ \mu = \mu(1);$$

so that (3) becomes

$$\mu(t) = \mu t \text{ where } 0 \leqq \mu \leqq \infty. \tag{4}$$

Furthermore we introduce the notation for $k \geqq 1$:

$$\begin{aligned} p_k(t) &= P\{N(t) = k\}, \\ r_k(t) &= P\{N(t) \geqq k\} = \sum_{j=k}^{\infty} p_j(t), \\ \lambda_k &= \lim_{t \downarrow 0} \frac{r_k(t)}{t}, \end{aligned}$$

whenever the limit exists. The process is said to be **regular** when $\lambda_2 = 0$.

The theorems by Khintchine, Dobrushin, and Korolyuk may be stated as follows (originally given for the strictly stationary case).

KHINTCHINE'S THEOREM. *λ_1 always exists: $0 \leqq \lambda_1 \leqq \infty$.*

DOBRUSHIN'S THEOREM. *If $\mu < \infty$ and there are no multiple points, then $\lambda_2 = 0$.*

KOROLYUK'S THEOREM. *If $\lambda_2 = 0$, then $\lambda_1 = \mu \leqq \infty$.*

It is the object of this note to formulate natural extensions of these results and give very simple proofs of them.

PROPOSITION 3. *If for some $k \geqq 1$ we have $\lambda_{k+1} = 0$, then*

$$\mu = \lim_{t \downarrow 0} \left(\sum_{j=1}^{k} \frac{r_j(t)}{t} \right). \tag{5}$$

For $k = 1$ this reduces to Korolyuk's theorem, which contains Khintchine's in the regular case. In general, the existence of λ_j for $2 \leqq j \leqq k$ is neither postulated nor implied. It is known (Khintchine [2], 3.8) that all λ_k exist for a strictly stationary process *without after effect*, that is, a **compound Poisson process** (see below).

PROPOSITION 4. *If all λ_k exist for $2 \leqq k < \infty$, finite or infinite, then*

$$\mu = \sum_{k=1}^{\infty} \lambda_k. \tag{6}$$

PROPOSITION 5. *Let $k \geqq 1$. If $\mu < \infty$ and there are no points with multiplicity $\geqq k+1$, then $\lambda_{k+1} = 0$. The converse is true for $\mu \leqq \infty$.*

For $k = 1$, the first part is Dobrushin's theorem; the second part is trivial for any k (cf. Cramér-Leadbetter [1], page 54).

PROPOSITION 6. *There is a strictly stationary counting process without any multiple point for which*

$$\mu = \lambda_1 = \infty, \quad 0 < \lambda_k < \infty \text{ for } k \geqq 2. \tag{7}$$

Further relevant facts will be mentioned at the end of section 3.

3. Proofs of the propositions. Let us begin by writing the elementary formula

$$E\{N(t)\} = \sum_{k=1}^{\infty} P\{N(t) \geqq k\}$$

in terms of our notation above:

$$\mu = \frac{\mu(t)}{t} = \sum_{k=1}^{\infty} \frac{r_k(t)}{t}. \tag{8}$$

We may thus regard the announced propositions as a study of the limiting form of the relation (8) as we let $t \downarrow 0$ and try to take the limits inside the summation—a meet game in analysis, made interesting here by the probabilistic interpretations.

Proof of Proposition 3. For each $t > 0$, we have (with an obvious abridging of notation)

$$N[0,1-t] \leqq \sum_{n=0}^{m-1} N[nt,(n+1)t], \qquad m = \left[\frac{1}{t}\right].$$

It is plain that for each integer $M > 0, \{N[0,1] \leqq M\} \subset \{N[nt,(n+1)t] \leqq M\}$ for $0 \leqq n \leqq m-1$. Hence we have

$$\begin{aligned}\int_{N[0,1]\leqq M} N[0,1-t]\,dP &\leqq \sum_{n=0}^{m-1} \int_{N[nt,(n+1)t]\leqq M} N[nt,(n+1)t]\,dP \\ &= m \int_{N[0,t]\leqq M} N[0,t]\,dP\end{aligned}$$

by crude stationarity in the last equation. From the definitions of the quantities involved, we have

$$\int_{N[0,t]\leqq M} N[0,t]\,dP = \sum_{j=1}^{M} j\,p_j(t) \leqq \sum_{j=1}^{M} r_j(t).$$

It follows that

$$\int_{N[0,1]\leqq M} N[0,1-t]\,dP \leqq \frac{t+1}{t} \sum_{j=1}^{M} \frac{r_j(t)}{t}. \tag{9}$$

Letting $t\downarrow 0$ and using the monotone convergence theorem on the left together with Proposition 1, we obtain

$$\int_{N[0,1]\leqq M} N[0,1]\,dP \leqq \underline{\lim}_{t\downarrow 0} \sum_{j=1}^{k} \frac{r_j(t)}{t} \tag{10}$$

because the hypothesis $\lambda_{k+1} = 0$ forces $\lambda_j = 0$ for $j \geqq k+1$. Letting $M\uparrow\infty$ and observing that the reverse inequality for $\overline{\lim}_{t\downarrow 0}$ is trivial we get (5).

Proof of Proposition 4. It is plain from (8) (a case of Fatou's lemma) that

$$\mu \geqq \sum_{k=1}^{\infty} \lambda_k. \tag{11}$$

Letting $t\downarrow 0$ in (9) as before, then $M\uparrow\infty$, we obtain the reverse of (11) and so (6).

Proof of Proposition 5. Fix k and define first for each interval I:

$$\xi(I) = 1_{[N(I)\geqq k+1]}$$

and then for $t > 0$:

$$\eta(t) = \sum_{n=0}^{m-1} \xi[nt,(n+1)t], \qquad m = \left[\frac{1}{t}\right].$$

Thus $\eta(t)$ is the number of subintervals $[nt,(n+1)t]$ in which there are at least $k+1$ points counted with multiplicity. If no point has multiplicity $\geqq k+1$, then each $S(\omega)$ is a discrete set of "points with multiplicity $\leqq k$." If $\delta(\omega)$ denotes the minimum distance between the points of $S(\omega)\cap I$ without their multiplicities, then $\eta(t,\omega) = 0$ for $0 < t < \delta(\omega)$. Thus

$$P\left\{\lim_{t\downarrow 0} \eta(t) = 0\right\} = 1. \tag{12}$$

On the other hand, it is obvious that $\eta(t) \leqq N([0,1])$, where the right member above has expectation μ. If $\mu < \infty$ then by dominated convergence

$$\lim_{t\downarrow 0} E\{\eta(t)\} = 0.$$

Now we have by crude stationarity

(13) $$E\{\eta(t)\} = \left[\frac{1}{t}\right] E\{\xi[0,t]\} = \left[\frac{1}{t}\right] r_{k+1}(t).$$

Hence as $t \downarrow 0$ the last term tends to 0, i.e., $\lambda_{k+1} = 0$.

The converse will be shown in an extended form in Proposition 8 below.

Normally speaking, the condition $\lambda_k > 0$ should signal the existence of points with multiplicity $\geqq k$. This is the case for a compound Poisson process, which may be derived from a simple Poisson process by randomizing the multiplicity of each point according to a fixed distribution and independently over all the points. It is also the case, in a more general way, for a continuous time homogeneous Markov chain, where the situation is indicated by the formula, in standard notation:

$$\lim_{t \downarrow 0} \frac{p_{jk}(t)}{t} = q_{jk}, \qquad j \neq k.$$

Proposition 5 says that it is always true for a crudely stationary counting process when $\mu < \infty$. It may or may not be surprising that this is no longer so when $\mu = \infty$, as shown in the following counterexample.

Example (*proof of Proposition* 6). For each $\lambda > 0$ let $N^{(\lambda)}$ be a (simple) Poisson process on R with *intensity* $\lambda_1 = \lambda$; namely,

$$\mu^{(\lambda)}(t) = E\{N^{(\lambda)}[0,t]\} = \lambda t;$$

$$r_{k+1}^{(\lambda)}(t) = 1 - e^{-\lambda t} \sum_{j=0}^{k} \frac{(\lambda t)^j}{j!}, \qquad k \geqq 0.$$

Let N denote the counting process obtained by randomizing λ according to the distribution F. Specifically, we choose F to have the density f given below:

$$f(\lambda) = \begin{cases} \dfrac{1}{\lambda^2} & \text{if } \lambda \geqq 1, \\ 0 & \text{if } 0 < \lambda < 1. \end{cases}$$

Since each $N^{(\lambda)}$ is strictly stationary, so is N. Since no $N^{(\lambda)}$ has any multiple points, nor does N. We have

$$\mu = E\{N[0,1]\} = \int_0^\infty \lambda f(\lambda)\, d\lambda = \int_1^\infty \frac{1}{\lambda^2} d\lambda = +\infty;$$

$$\frac{r_{k+1}(t)}{t} = \int_1^\infty \left\{1 - e^{-\lambda t} \sum_{j=0}^{k} \frac{(\lambda t)^j}{j!}\right\} \frac{1}{t\lambda^2} d\lambda.$$

Making the change of variable $t\lambda = u$, we obtain

(14) $$\frac{r_{k+1}(t)}{t} = \int_t^\infty \left\{1 - e^{-u} \sum_{j=0}^{k} \frac{u^j}{j!}\right\} \frac{du}{u^2}.$$

As $t \downarrow 0$ the limit λ_{k+1} is therefore just the integral in (14) with $t = 0$. Thus it is clear that $0 < \lambda_{k+1} < \infty$ for $k \geqq 1$, but

$$\lambda_1 = \int_0^\infty \frac{1 - e^{-u}}{u^2} du = + \infty . \tag{15}$$

Perhaps the point of the theorems by Korolyuk and Dobrushin is the equality

$$\mu = \lambda_1 . \tag{16}$$

There is no reason to expect anything of the sort when multiple points are allowed, as in compound Poisson processes. Thus the two theorems together settle the case where $\mu < \infty$ and there are no multiple points. The case $\mu = \infty$ could be facilely dismissed by applied probabilists as "possessing no practical interest" (see N. B. at the end of this paper). Nevertheless, let us point out that as a corollary to Propositions 4 and 5, (16) is also true when $\mu = \infty$, and for some $k \geqq 2$ there are no points of multiplicity $\geqq k$. For then by (5) and Khintchine's theorem we have

$$\infty = \lambda_1 + \lim_{t \downarrow 0} \sum_{j=2}^{k} \frac{r_j(t)}{t} \tag{17}$$

and the last limit must be finite if $\lambda_1 < \infty$, since r_j decreases as j increases. Thus $\lambda_1 = \infty = \mu$. Another case where this is so is given in the example above. Leadbetter [3] has shown that if there are no multiple points, then $\mu = \infty$ implies $\lambda_1 = \infty$.

3. Extension to several dimensions. We turn now to the consideration of theorems of the above type when R is replaced by R^d, the Euclidean space of dimension d. There are recent studies of point processes in which the points belong to a more general topological space, but so far as I am aware these are not relevant to the questions at hand. The extensions to R^d decidedly possess practical interest, since scientists do count particles with a grid under the microscope, etc. As no new difficulty arises when $d \geqq 3$, we shall take $d = 2$.

Call I an **interval** in R^2 iff it is a bounded parallelogram with its sides parallel to the coordinate axes, but may or may not include all its boundary. Denote its side lengths by $a(I)$ and $b(I)$, its area and diameter by

$$|I| = a(I)b(I), \qquad d(I) = \sqrt{a(I)^2 + b(I)^2},$$

and put

$$\rho(I) = \frac{b(I)}{a(I)} .$$

For each ρ, where $0 < \rho < \infty$, the family of such intervals with $\rho(I) = \rho$ will be denoted by $\mathscr{K}(\rho)$, for example, when $\rho = 1$ these are squares. The family of all intervals will be denoted by $\mathscr{K} = \bigcup_{0<\rho<\infty} \mathscr{K}(\rho)$.

Under assumptions analogous: o (A) and (B), the process $\{N(I,\omega)\}$ with $I \in \mathscr{K}$, $\omega \in \Omega$, will be called a crudely stationary counting process on R^2 iff whenever I and J are two closed intervals of the same area, the random variables $N(I)$ and $N(J)$ have the same distribution. Analogues of Propositions 1 and 2 then hold, but be careful: the intervals I_n in the analogue of (2) must be assumed to be uniformly bounded, in other words, contained in a fixed interval. We have as the analogue to (4), for each interval I:

$$E\{N(I)\} = \mu|I|, \quad \text{where } \mu = E\{N(Q)\}, \tag{18}$$

and Q is a unit square in $\mathscr{K}$.

Fix ρ and a member J of $\mathscr{K}(\rho)$. Then all members of $\mathscr{K}(\rho)$ are congruent to tJ for some $t > 0$, where tJ is an interval homothetic to J at the ratio $t:1$, so that $|tJ| = t^2|J|$. If we now restrict ourselves to members of $\mathscr{K}(\rho)$, we may put

$$r_k(t) = P\{N(tJ) \geqq k\},$$

$$\lambda_k(\rho) = \lim_{t \downarrow 0} \frac{r_k(t)}{t^2|J|},$$

whenever the limit exists. Then Khintchine's theorem as well as Propositions 3, 4 and 5 can all be extended to this case. For instance, we have the following trivial extension of the well-known subadditivity lemma used by Khintchine.

LEMMA. *Let ϕ on $[0,\infty)$ be non-negative and have the following property: whenever $0 < t \leqq ns$, where n is a positive integer, we have*

$$\phi(t) \leqq n^2\phi(s).$$

Then we have

$$\lim_{t \downarrow 0} \frac{\phi(t)}{t^2} = \sup_{t>0} \frac{\phi(t)}{t^2} \leqq \infty.$$

If we set $\phi(t) = r_1(t)$, then Boole's inequality and crude stationarity imply that ϕ satisfies the conditions of the lemma, from which the extended Khintchine theorem follows. Similarly, the proofs of the other propositions carry over to the present case without any difficulty.

However, it is more interesting to consider the larger family $\mathscr{K}$ of all intervals. We then define for $k \geqq 1$:

$$\lambda_k = \lim_{\substack{d(I) \to 0 \\ I \in \mathscr{K}}} \frac{P\{N(I) \geqq k\}}{|I|} \tag{19}$$

whenever the limit exists. The methods used above can be modified to prove the cited propositions in the new context. Everything depends on the following elementary covering lemma:

LEMMA. *Let $I \in \mathscr{K}$ and $\varepsilon > 0$ be given; there exists $\delta = \delta(I, \varepsilon) > 0$ with the following property: For any $J \in \mathscr{K}$ with $d(J) < \delta$, we can find J_j, $1 \leqq j \leqq l$, which are disjoint (apart from sides) and all congruent to J, and which satisfy*

$$I \subset \sum_{j=1}^{l} J_j \subset (1+\varepsilon)I.$$

The proof is omitted as geometrically obvious.

We now state and prove the theorems by Khintchine, Dobrushin and Korolyuk, leaving the previous extensions of the last two theorems to the reader.

PROPOSITION 7. *The limit λ_1 always exists, $\leqq \infty$.*

Proof. Denote by λ_1' the lower limit on the right side of (19), when $k = 1$. Let J_n be a sequence of intervals achieving this lower limit; thus

$$\lambda_1' = \lim_{n \to \infty} \frac{P\{|J_n| \geqq 1\}}{|J_n|}.$$

Since $d(J_n) \to 0$, we may apply the covering lemma to I and J_n for all n such that $d(J_n) < \delta$. Thus we have J_{nj}, $1 \leqq j \leqq l_n$, all congruent to J_n such that

$$I \subset \bigcup_{j=1}^{l_n} J_{nj} \subset (1+\varepsilon)I. \tag{20}$$

It follows from the first inclusion and Boole's inequality that

$$\{N(I) \geqq 1\} \subset \bigcup_{j=1}^{l_n} \{N(J_{nj}) \geqq 1\};$$

and consequently by crude stationarity

$$P\{N(I) \geqq 1\} \leqq l_n P\{N(J_n) \geqq 1\}.$$

On the other hand, the second inclusion in (20) implies that

$$l_n |J_n| \leqq (1+\varepsilon)^2 |I|.$$

Combining the last two inequalities, we obtain

$$\frac{P\{N(I) \geqq 1\}}{|I|} \leqq (1+\varepsilon)^2 \frac{P\{N(J_n) \geqq 1\}}{|J_n|}. \tag{21}$$

Letting $n \to \infty$ in (21) and then $\varepsilon \to 0$, we see that the left member of (21) does not exceed λ_1'. Since I is arbitrary, this means

$$\sup_{I \in \mathscr{K}} \frac{P\{N(I) \geqq 1\}}{|I|} \leqq \lambda_1';$$

the more so if the "sup" above is replaced by the upper limit as $d(I) \to 0$. Therefore λ_1 exists.

PROPOSITION 8. *If there are no multiple points and $\mu < \infty$, then $\lambda_2 = 0$. If $\lambda_2(\rho) = 0$ for some $\rho > 0$, then almost surely there are no multiple points.*

Proof. For every J in $\mathscr{K}$, we put

$$\xi(J) = I_{[N(J) \geqq 2]}.$$

Let I and J_n be given in $\mathscr{K}$, where $d(J_n) \to 0$; as in the preceding proof, we have (20) for large n. Now define

$$\eta_n(I) = \sum_{j=1}^{l_n} \xi(J_{nj}).$$

If there are no multiple points, then just as in the proof of Proposition 3,

$$P\left\{\lim_{n\to\infty} \eta_n(I) = 0\right\} = 1.$$

Since $\eta_n(I) \leqq N((1+\varepsilon)I)$ there is dominated convergence so that

$$\lim_{n\to\infty} E\{\eta_n(I)\} = 0.$$

But from (20) and crude stationarity

$$E\{\eta_n(I)\} \geqq l_n E\{\xi(J_n)\} \geqq |I| \frac{P\{N(J_n) \geqq 2\}}{|J_n|}.$$

Hence

$$\lim_{n\to\infty} \frac{P\{N(J_n) \geqq 2\}}{|J_n|} = 0.$$

This being true for any sequence J_n in $\mathscr{K}$ with $d(J_n) \to 0$, we have $\lambda_2 = 0$.

Conversely, suppose $\lambda_2(\rho) = 0$. Choose any J from $\mathscr{K}(\rho)$ and divide it into 4^n disjoint J_{nj} all congruent to $2^{-n}J$. (This is nothing but Weierstrass' bisection argument.) Clearly

$$\{\xi(J) > 0\} \subset \bigcup_{j=1}^{4^n} \{\xi(J_{nj}) > 0\};$$

hence

$$\begin{aligned} P\{N(J) \geqq 2\} &= P\{\xi(J) > 0\} \leqq 4^n P\{\xi(2^{-n}J) > 0\} \\ &= |J| \frac{P\{N(2^{-n}J) \geqq 2\}}{|2^{-n}J|}. \end{aligned}$$

The last term tends to zero by hypothesis, and J is arbitrary; it follows that there is no multiple point (with probability one).

PROPOSITION 9. *If* $\lambda_2(\rho) = 0$ *for some* $\rho > 0$, *then* $\lambda_1 = \mu \leqq \infty$.

Proof. We have remarked that the extension of Korolyuk's theorem is easy for the family $\mathscr{K}(\rho)$. Hence if $\lambda_2(\rho) = 0$ then $\lambda_1(\rho) = \mu$. But by proposition 7, $\lambda_1(\rho) = \lambda_1$ for every ρ.

In conclusion, we may ask what family of figures satisfies a covering property as stated in the Lemma, or some weaker form of it which will still serve the purpose. If we confine ourselves to polygonal ones, then one family is that of all such figures which can be used to *pave* the plane, such as triangles and honeycomb-like hexagons (not necessarily regular) as well as our family $\mathscr{K}$. Paving figures with curved boundaries may be considered provided the boundaries are smooth enough. On the other hand, disks seem to be out, despite Vitali's covering theorem. Nevertheless, are there appropriate extensions of the results discussed here to such figures as disks?

N.B. It is not a mere flight of rhetoric to say that in many mathematical questions, one must ponder over the infinite in order fully to comprehend the finite. Surely the most celebrated instance of this in the history of probability is the St. Petersburg Paradox dealing with the law of large numbers when the mathematical expectation is infinite. A similar situation is the central limit theorem under Lindeberg's condition, when the variance is infinite. Perhaps more relevant to the subject of this note is the existence of quasi-stationary distribution in a recurrent Markov chain, when the steady state must be described by an infinite total mass. This plays a basic role in the deeper parts of the theory. The possibility of infinitely many jumps in finite time, corresponding to the case where $P\{N(t) = +\infty\} > 0$, in the notation of this note is the origin of modern boundary theory, which ought to find applications in various explosive or rapidly changing phenomena. Applied mathematicians are all too apt to dismiss a somewhat delicate situation as pathological or impractical simply because their tools are too crude to cope with them, and then justify this on spurious grounds. It is by no means clear that Nature operates on finiteness assumptions, otherwise why are there infinitely many primes?

Added in proof: I am indebted to Daley and Vere-Jones for the remark that in R^1, $P\{1 \leqq N(t) \leqq k\}$ is subadditive in t, hence λ_k exists for $k \geqq 2$ provided $\lambda_1 < \infty$. A similar result holds for the λ_k defined in (19) by a simple modification of the proof of Proposition 7. See also a forthcoming paper by R. K. Milne in *The Annals of Mathematical Statistics*.

Research supported in part by the Air Force Office of Scientific Research, Air Force Systems Command, USAF under AFOSR Contract F44620-67-C-0049.

References

1. H. Cramér and M. R. Leadbetter, Stationary and Related Stochastic Processes, Wiley, New York, 1967.

2. Y. A. Khintchine, Mathematical Methods in the Theory of Queueing, Griffin, London, 1960.

3. M. R. Leadbetter, On three basic results in the theory of stationary point processes, Proc. Amer. Math. Soc., 19 (1968) 115–117.

Istituto Nazionale di Alta Matematica
Symposia Mathematica
Volume IX (1972)

ON THE FUNDAMENTAL HYPOTHESES OF HUNT PROCESSES (*) (**)

KAI LAI CHUNG

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $X = \{X_t, t > 0\}$ a homogeneous Markov process adapted to the Borel subfields $\{\mathcal{F}_t\}$ of $\mathcal{F}$. The process X is supposed to be measurable and it takes values in $(\boldsymbol{E}, \mathcal{E})$, where $\boldsymbol{E}$ is a locally compact space with countable base and $\mathcal{E}$ its Borel field. The transition function $P_t(x, A)$, $t > 0$, $x \in \boldsymbol{E}$, $A \in \mathcal{E}$, is Borelian in (t, x). The field family $\{\mathcal{F}_t\}$ is augmented by all P-null sets but not necessarily right continuous. A stopping time relative to $\{\mathcal{F}_{t+}\}$ will be called « optional » here. The augmented Borel field generated by (...) will be denoted by $\sigma(...)$. C_K denotes the class of functions continuous on $\boldsymbol{E}$ and having compact supports. « Almost surely » (« a.s. ») refers to P and is often omitted in an obvious context. Any statement below regarding X_T is automatically understood to be under the proviso « a.s. on $\{T < \infty\}$ » since X_∞ is not defined. « The path » is a circumlocution for « almost all paths ». The symbol « $\uparrow$ » means « increasing » but not necessarily strictly. Other terminology, notation and conventions follow generally those in [1] and [5].

The three fundamental hypotheses of a Hunt process are as follows:

(a) the path $t \to X(t, \omega)$ is right continuous and has left limits everywhere in $(0, \infty)$.

(b) For any sequence of optional times $\{T_n\}$ such that $T_n \uparrow T$, we have

$$\lim_n X_{T_n} = X_T \,. \tag{1}$$

This property will be called « quasi left continuity ».

(*) I risultati conseguiti in questo lavoro sono stati esposti nella conferenza tenuta il 16 marzo 1971.

(**) Research supported in part by the Air Force Office of Scientific Research, Air Force Systems Command. USAF under AFOSR Contract F44620-67-C-0049.

(c) For every optional T, $t>0$ and $f\in C_K$, we have

$$E\{f\circ X_{T+t}|\mathcal{F}_{T+}\}=P_tf\circ X_{T+}\,. \tag{2}$$

This property will be called the « strong Markov property », although it is given here in its tactical form. Note that $X_T\in\mathcal{F}_{T+}$ for every optional T, provided that X is progressively measurable and that is implied by right continuity of the path.

These hypotheses are finely meshed. For instance, right continuity, (b) and (c) together imply the existence of left limits in (a). The form of (c) is geared to the right continuity and we have underlined this by writing X_{T+} for X_T in the right member of (2). Speaking figuratively, the forces of the Hunt hypotheses are basically biased toward the right, protecting the left flank only by the rather heavy-handed (b). This is not unnatural because time flows in one direction and a handle on the future is worth more than one on the past. The essential dissymmetry becomes manifest if one compares the « optionally measurable » (= « *bien mesurable* ») with the « foreseeably measurable » (= « *prévisible* »). The former, generated by right continuous adapted processes, encompasses the latter, generated by left continuous ones. However, the situation is reversed by a reversing of time. It is found in [3] that one cannot keep the reverse process natural, namely left continuous, and yet fit it into a right-oriented mold such as (c). This has to be replaced by a left-oriented counterpart which is christened the « moderate Markov property ». We shall examine here the possibility of postulating such a property for classes of homogeneous Markov process including both Hunt processes and their reverses; and its relation to quasi left continuity. We shall also pinpoint the difference between the left and right fields involved in the formulations.

We begin by recalling the definition of a « foreseeable » [or « predictable »] time; this was introduced by Meyer [6], see also Dellacherie [4].

DEFINITION 1: T (>0) is foreseeable iff there exists a sequence of optional times $\{T_n\}$ such that

$$\forall n\colon T_n<T \text{ and } T_n\uparrow T \text{ on } \Omega\,. \tag{3}$$

REMARK. If $\{T_n\}$ satisfies (3) on $\{T<\infty\}$, then $\{T_n\wedge n\}$ satisfies (3) as stated. The conditions may be qualified by « a.s. » which we have promised to omit. The possibility of $T=0$ may be included by a convention which we ignore.

The following lemma is fundamental. A ramified version was given in [3] as Lemma 6.1 without proof, but we have since realized that it is not quite so obvious.

LEMMA 1. If T is foreseeable, then there exists a sequence of countably-valued optional times $\{T_n\}$ satisfying (3).

PROOF. Let T_n be as in Definition 1, not necessarily countably-valued. For each n and k we put

$$T_n^{(k)} = \frac{[2^k T_n + 1]}{2^k}$$

so that $T_n^{(k)} > T_n$ and

$$\forall n: \lim_k T_n^{(k)} = T_n < T .$$

Hence there exists k_n such that

$$P\{T_n^{(k_n)} < T\} \geqslant 1 - \frac{1}{2^n}.$$

By the lemma of Borel-Cantelli, for almost every ω, there exists $N(\omega) < \infty$ such that

$$T_n^{(k_n)}(\omega) < T(\omega) , \qquad \forall n \geqslant N(\omega): \tag{4}$$

Put

$$S_m(\omega) = \inf_{n \geqslant m} T_n^{(k_n)}(\omega) . \tag{5}$$

Then almost surely

$$\inf_{n \geqslant m} T_n \leqslant S_m < T$$

by (4) and so $S_m \uparrow T$. Since

$$\liminf_n T_n^{(k_n)} \geqslant \lim_n T_n = T > S_m ,$$

the infimum in (5) is attained at a finite value of n and consequently S_m is countably-valued as each T_n is. Thus the sequence $\{S_n\}$ satisfies the requirements of the Lemma.

REMARK. Given any countable dense set D in $(0, \infty)$ we can require that the T_n in Lemma 1 to be D-valued. The preceding proof for

$$D = \left\{ \frac{j}{2^k} \middle| k \geqslant 1, \ j \geqslant 1 \right\}$$

needs only a technical modification to apply to any such D.

DEFINITION 2: Suppose that the path has left limits everywhere in $(0, \infty)$. Then X is said to have the moderate Markov property at a foreseeable T iff for every $t > 0$ and $f \in C_K$, we have

$$E\{f \circ X_{T+t} | \mathcal{F}_{T-}\} = P_t f \circ X_{T-} . \tag{6}$$

X is said to have the property iff this is true for every foreseeable T.

Recall that $\mathcal{F}_{T-}$ for *any* random variable $T > 0$ is the Borel subfield of $\mathcal{F}$ generated by sets of the form $\{T > t\} \cap \Lambda_t$ where t ranges over $(0, \infty)$ and for each t, Λ_t ranges over $\mathcal{F}_t$. Note that $\mathcal{F}_{T-}$ is augmented as each $\mathcal{F}_t$ is, $\mathcal{F}_{T-} \in \mathcal{F}_{T+}$ for optional T, and $X_{T-} \in \mathcal{F}_{T-}$. [This is easy for foreseeable T, for extension see Lemma 2 below]. An equivalent form of Definition 2 under further conditions will be given in Theorem 3.

Let U^α be the α-potential of (P_t), namely for every bounded or positive $f \in \mathcal{E}$:

$$U^\alpha f = \int_0^\infty \exp[-\alpha t] P_t f \, dt .$$

THEOREM 1. Assume either condition (a) above, or the following condition:

(d) the path is left continuous in $(0, \infty)$.

Then X has the moderate Markov property at T if for every $\alpha > 0$ and $f \in C_K$ there exists a sequence of countably-valued optional $\{T_n\}$ satisfying (3) such that a.s. on $\{T < \infty\}$ we have

$$\lim_n U^\alpha f \circ X_{T_n} = U^\alpha f \circ X_{T-} . \tag{7}$$

Conversely if this is true then (7) holds for every such sequence.

PROOF. Note that the limit above exists a.s. because $n \to U^\alpha f \circ X_{T_n}$ is a bounded supermartingale. To prove that (7) implies the moderate Markov property at T, let $\{T_n\}$ be as asserted. Since T_n is countably-valued, the strong Markov property is applicable at T_n without the « + » in T_n and yields

$$\exp[-\alpha T_n] U^\alpha f \circ X_{T_n} = E\left\{ \int_{T_n}^\infty \exp[-\alpha t] f \circ X_t \, dt \middle| \mathcal{F}_{T_n} \right\}. \tag{8}$$

Letting $n \to \infty$ on $\{T < \infty\}$, so that $\mathcal{F}_{T_n} \uparrow \mathcal{F}_{T-}$ while the integral on the right side converges to that over (T, ∞), we obtain by an easy

dominated convergence theorem for conditional expectations (erstwhile cited as Hunt's lemma but given also by Blackwell and Dubins) and using (7):

$$\exp[-\alpha T]\, U^\alpha f \circ X_{T-} = E\left\{ \int_T^\infty \exp[-\alpha t] f \circ X_t \, dt \Big| \mathcal{F}_{T-} \right\}$$

$$= \exp[-\alpha T]\, E\left\{ \int_0^\infty \exp[-\alpha t] f \circ X_{T+t} \, dt \Big| \mathcal{F}_{T-} \right\}.$$

It follows that for every $\Lambda \in \mathcal{F}_{T-}$, we have

$$\int_\Lambda U^\alpha f \circ X_{T-} \, dP = \int_\Lambda \left[\int_0^\infty \exp[-\alpha t] f \circ X_{T+t} \, dt \right] dP ,$$

and consequently by Fubini's theorem:

$$(9) \qquad \int_0^\infty \exp[-\alpha t] dt \left[\int_\Lambda P_t f \circ X_{T-} \, dP \right] = \int_0^\infty \exp[-\alpha t] dt \left[\int_\Lambda f \circ X_{T+t} \, dP \right].$$

This being true for all $\alpha > 0$, it follows from the uniqueness theorem for Laplace transforms that we have for almost all t with respect to Lebesgue measure:

$$(10) \qquad \int_\Lambda P_t f \circ X_{T-} \, dP = \int_\Lambda f \circ X_{T+t} \, dP .$$

Now as functions of t, both members in (10) are right continuous under condition (a) and left continuous under condition (d). Therefore (10) holds for all $T > 0$. Since Λ is arbitrary, this is equivalent to (6).

Conversely, suppose (6) is true. Integrating it after multiplying through by $\exp[-\alpha t]$, we obtain after a change of variables:

$$(11) \qquad \exp[-\alpha T]\, U^\alpha f \circ X_{T-} = E\left\{ \int_T^\infty \exp[-\alpha t] f \circ X_t \, dt \Big| \mathcal{F}_{T-} \right\} .$$

Let $\{T_n\}$ be any countably-valued optional times satisfying (3) so that (8) holds as before. If $n \to \infty$ in (8) then its right member converges to that of (11), and (7) follows.

It is not difficult to deduce from Lemma 1, Theorem 1 and Meyer's capacity theorem for a foreseeably measurable set the following result, which should be compared with T. 15 of Chapter XIV in [5].

COROLLARY: If X has left continuous paths and the moderate Markov property, then for every $\alpha > 0$, and bounded $f \in \mathcal{E}$, the function

$$t \to U^{\alpha} f \circ X_t \tag{12}$$

is left continuous in $(0, \infty)$ almost surely.

We shall strengthen condition (a) a little by adding the following and referring to their conjunction as (a'):

$$\forall x \in E,\ f \in C_E:\ \lim_{t \downarrow 0} P_t f(x) = f(x)\,. \tag{13}$$

This is related to «normality» as defined in [1] and the effect is to exclude the existence of «branch points».

THEOREM 2. Under condition (a'), if X has the moderate Markov property, then it is quasi left continuous.

PROOF. First suppose that $\{T_n\}$ is a sequence of optional times satisfying (3) so that T is foreseeable. By hypothesis, (6) holds for every $t > 0$. Letting $t \downarrow 0$ and using (13) on the right side of (6), we obtain

$$E\{f \circ X_{T+} | \mathcal{F}_{T-}\} = X_{T-}\,. \tag{14}$$

This being true for every $f \in C_E$, it follows by a well-known lemma that

$$X_{T+} = X_{T-} = \lim_n X_{T_n+}\,. \tag{15}$$

It remains to remove the condition of strict inequality in (3). This is easy but should be done carefully. Suppose then $T_n \uparrow T$, and set

$$\Lambda = \{\forall n:\ T_n < T < \infty\} \in \mathcal{F}_{T-};$$

$$T' = \begin{cases} T & \text{on } \Lambda\,, \\ \infty & \text{on } \Lambda^c; \end{cases}$$

$$T'_n = \begin{cases} T_n & \text{on } \{T_n < T\} \in \mathcal{F}_{T_n+}\,, \\ \infty & \text{on } \{T_n = T\}\,. \end{cases}$$

Then each T'_n is optional. On $\Lambda = \{T' < \infty\}$, we have

$$(16) \qquad T_n = T'_n < T' = T\,, \qquad T'_n \uparrow T;$$

hence T' is foreseeable by the remark after Definition 1. Thus by (15) applied to T'_n and T:

$$X_{T'+} = X_{T'-} = \lim_n X_{T'_n+}$$

and consequently by the first part of (16):

$$X_{T+} = \lim_n X_{T_n+}$$

on Λ, while this is trivial on $\Lambda^c \cap \{T < \infty\}$. This establishes the quasi left continuity.

We use the notation $\mathcal{F}'_T = \sigma(X_{T+t}, t > 0)$.

THEOREM 3: Under condition (a') or (d) the moderate Markov property is equivalent to the following proposition: for every foreseeable T, $\Lambda \in \mathcal{F}_{T-}$ and $M \in \mathcal{F}'_T$, we have

$$(17) \qquad P\{\Lambda \cap M | X_{T-}\} = P\{\Lambda | X_{T-}\}\, P\{M | X_{T-}\}\,.$$

PROOF. We omit the familiar proof modeled after that for the usual strong Markov property, except to make the crucial remark that under (a') or (d), we have for each $t \geq 0$, almost surely

$$X_{T+t-} = X_{T+t}\,.$$

This follows from Theorem 2 under (a') and from the hypothesis under (d).

Let us remark that if we define the strong Markov property as in (c) above, *separately* from right continuity of the path, then its extension to a form like (17) for an optional T and with the $\mathcal{F}_{T-}$, X_{T-} there replaced by $\mathcal{F}_{T+}$, X_{T+}, will need some further condition too.

Despite the connotation of the adjectives employed, the *moderate* Markov property is not a consequence of the *strong* one. The following case is rather trivial but it includes at least a Hunt process. Note that taken separately from (a), the existence of X_{T+} is part of the tacit assumption in (c).

THEOREM 4. Under (a'), strong Markov property and quasi left continuity together imply the moderate Markov property.

PROOF: Let T be foreseeable and $\{T_n\}$ as in (3). By quasi left continuity we have

$$\lim_n X_{T_n+} = X_{T+}\,. \tag{18}$$

$$\forall t > 0:\ P_t f \circ X_{T_n+} \in \mathcal{F}_{T_n+} \subset \mathcal{F}_{T-}\,.$$

Letting $t \downarrow 0$ and using (13) we obtain $X_{T_n+} \in \mathcal{F}_{T-}$, hence $X_{T+} \in \mathcal{F}_{T-}$ by (18). Now if we condition (2) with respect to $\mathcal{F}_{T-}$ and use (18) again, we obtain

$$E\{f \circ X_{T+t} | \mathcal{F}_{T-}\} = P_t f \circ X_{T+} = P_t f \circ X_{T-}$$

which is (2).

The switching from $T-$ to $T+$ makes the difference in the two forms of Markov property, and this will be expressed by the field equation (22) below. This kind of question was treated by Meyer as « times of discontinuity » of the field family $\{\mathcal{F}_t\}$, before the notion of $\mathcal{F}_{T-}$ was introduced in [2]. We think the present formulation making use of this notion is more perspicuous than the usual one (cf. Proposition IV.4.1 and IV.4.2 in [1]). The following lemma is needed only in the special case of an optional T but is given in its general form to illustrate the possibilities, now that the theory of Markov processes has progressed sufficiently to take serious account of non-optional times such as « last exit times » and « starting times ».

In the lemma and its corollary below $\{X_t, \mathcal{F}_t\}$ is as before.

LEMMA 2: Suppose that $\{Y_t, t > 0\}$ is adapted to $\{\mathcal{F}_t\}$ and that the path $t \to Y(t, \omega)$ has left limits everywhere in $(0, \infty)$. Then for every random variable T, $0 < T < \infty$, we have

$$Y_{T-} \in \mathcal{F}_{T-}\,. \tag{19}$$

PROOF. It follows from the definition of $\mathcal{F}_{T-}$ that for any Borel set $B \subset (t, \infty)$ and any $\Lambda \in \mathcal{F}_t$, we have

$$\{T \in B\} \cap \Lambda \in \mathcal{F}_{T-}\,. \tag{20}$$

Put for every $n \geqslant 1$:

$$T_n = \frac{[2^n T - 1]}{2^n}$$

so that $T_n < T$, $T_n \uparrow T$. We have then for any $\Lambda \in \mathcal{F}_{m/2^n}$,

$$\left\{T_n = \frac{m}{2^n}\right\} \cap \Lambda = \left\{\frac{m+1}{2^n} \leqslant T < \frac{m+2}{2^n}\right\} \cap \Lambda \in \mathcal{F}_{T-}.$$

by (20). Hence for any $A \in \mathcal{E}$:

$$\left\{T_n = \frac{m}{2^n};\ Y(T_n) \in A\right\} = \left\{T_n = \frac{m}{2^n};\ Y\left(\frac{m}{2^n}\right) \in A\right\} \in \mathcal{F}_{T-}$$

since Y is adapted. It follows that $Y(T_n) \in \mathcal{F}_{T-}$ and (19) follows since $Y(T_n) \to Y(T-)$.

COROLLARY. Let X be progressively measurable, and $\varphi \in \mathcal{E}$ be such that $t \to \varphi(X_t)$ is integrable in $(0, \infty)$. Then for any random variable T, $0 < T < \infty$, we have

$$\int_0^T \varphi(X_t)\, dt \in \mathcal{F}_{T-}. \tag{21}$$

PROOF. Apply the lemma to

$$Y_t = \int_0^t \varphi(X_s)\, ds.$$

Then $Y_t \in \mathcal{F}_t$ by progressive measurability and $t \to Y(t, \omega)$ is even continuous. The case $T = 0$ or $T = \infty$ is a matter of convention.

In the Theorem below we take $\mathcal{F}_\infty = \sigma(X_t, 0 \leqslant t < \infty)$.

THEOREM 5. If X is right continuous and has the strong Markov property given in (c), then for every optional T we have

$$\mathcal{F}_{T+} = \mathcal{F}_{T-} \vee \sigma(X_{T+}) \tag{22}$$

[= the Borel field generated by $\mathcal{F}_{T-}$ and X_{T+}].

PROOF. Except for Lemma 2 the proof can be lifted out of pp. 171-172 of [1], and here is the gist. The field $\mathcal{F}_\infty$ is generated by sets H of the form (where $\alpha_k > 0$, $f_k \in C_K$, $l < \infty$):

$$\prod_{k=1}^{l} \int_0^\infty \exp[-\alpha t] f \circ X_t\, dt =$$

$$= \prod_i \int_0^T \exp[-\alpha_i t] f_i \circ X_t\, dt \prod_j \int_T^\infty \exp[-\alpha_j t] f_j \circ X_t\, dt.$$

The first product on the right side belongs to $\mathcal{F}_{T-}$ by Lemma 2; the second conditioned on $\mathcal{F}_{T+}$ belongs to $\sigma(T, X_{T+})$ by the Strong Markov property. Since $T \in \mathcal{F}_{T-}$ it follows that

$$E\{H|\mathcal{F}_{T+}\} \in \mathcal{F}_{T-} \vee \sigma(X_{T+})$$

and this ends the proof quickly.

COROLLARY: Under conditions (a), (b) and (c): « accessible » = = « foreseeable ».

PROOF: Equation (22) and T. II.46 of [4].

Added in proof. – A simpler proof of Theorem 5 is given in my note « Some universal field equations », to appear in *Séminaire de probabilités* VI, Université de Strasbourg.

Testo pervenuto il 16 marzo 1971.
Bozze licenziate il 10 aprile 1972.

REFERENCES

[1] R. M. BLUMENTHAL and R. K. GETOOR, *Markov processes and potential theory*, Academic Press (1968).

[2] K. L. CHUNG and J. L. DOOB, *Fields, optionality, and measurability*, Amer. J. Math., 87 (1964), 397-424.

[3] K. L. CHUNG and J. B. WALSH, *To reverse a Markov process*, Acta Math., 123 (1969), 225-251.

[4] C. DELLACHERIE, *La théorie générale des processus*, Séminaire de probabilités, Université de Strasbourg (1969).

[5] P. A. MEYER, *Processus de Markov*, Springer-Verlag (1967).

[6] P. A. MEYER, *Guide détaillé de la théorie « générale » des processus*, Séminaire de probabilités II, Université de Strasbourg, Springer-Verlag (1968).

Reprinted with kind permission.

Ann. Inst. Fourier, Grenoble
23,3-(1973), 313-322

PROBABILISTIC APPROACH IN POTENTIAL THEORY TO THE EQUILIBRIUM PROBLEM

by Kai Lai CHUNG (1)

The problem of equilibrium is the first problem for the ancient period of potential theory recounted by Brelot in his recent historical article [1]. The existence of an equilibrium measure for the Newtonian potential was affirmed by Gauss as the charge distribution on the surface of a conductor which minimizes the electrostatic energy. But it was Frostman who rigorously established the existence in his noted thesis (1935), and extended it to the case of M. Riesz potentials. Somewhat earlier, F. Riesz had given his well-known decomposition for a subharmonic function which is a closely related result. For further history and references to the literature see Brelot's article. From the viewpoint of probability theory, the equilibrium problem in the Newtonian case takes the following from :

$$\underline{\underline{P}}^x\{T_B < \infty\} = \int_{\partial B} u(x, y)\, \mu_B(dy). \tag{1}$$

Here the underlying process is a Brownian motion $\{X_t, t \geqslant 0\}$ in R^3 ; $\underline{\underline{P}}^x\{\cdots\}$ denotes the probability (Wiener measure) when all paths issue from the point x ; B is a compact set (the conductor body) ; $T_B = T_B(\omega)$ is the hitting time of B by the path ω :

$$T_B(\omega) = \inf\{t > 0 \mid X_t(\omega) \in B\} ;$$

∂B is the boundary of B ; $u(x, y)$ is the associated potential density

$$u(x, y) = \frac{1}{2\pi |x - y|} ;$$

(1) Research supported by the Office of Scientific Research of the United States Air Force, under AFOSR Contract F44620-67-C-0049.

and finally μ_B is the said equilibrium measure. Standard terminology and notation from the modern theory of Markov processes are used above. In the general setting to be considered here, the underlying process is a temporally homogeneous Markov process taking values in a topological space E which is locally compact and has a countable base with its Borel field $\mathcal{E}$. The transition semigroup will be assumed to be Borelian. However, we need not suppose the process to be a Hunt process ; in fact, strong Markov property will be used only peripherally toward the end, and quasi left continuity not at all. It is sufficient to assume that all paths are right continuous and have left limits in the time interval $[0, \infty)$. No overt duality assumptions are made in establishing the general formula (17) below.

A probabilistic proof of (1) is given in Ito-McKean [2, pp. 248ff.] which leans heavily on special analytic properties of the Brownian motion semigroup. In another paper, McKean [3] discussed a probabilistic interpretation of the result in a more general case, bringing in a number of things (capacity, Ueno's result, Weyl's lemma, etc.) which seem to obscure the real issue and leave the upshot unclear. For some reason the notion of a last exit time, which is manifestly involved in the arguments, would not be dealt with openly and directly. This may be partially due to the fact that such a time is not an "optional" (or "stopping") time, does not belong to the standard equipment, and so must be evaded at all costs. Actually the notion has been introduced to great advantage in Markov chains and the associated boundary theory, although it was only during the last few years that it became formalized (with considerable loss of intuition) under the name "co-optional". In the present approach it turns out to be a tame thing and leads very quickly to the classical results of Gauss-M. Riesz-Frostman, without any unnecessary complications. Moreover, pursuance of this simple idea yields a more complete solution of the equilibrium problem for a broad class of Markov processes. A historical note may be added here : a probabilistic solution to Dirichlet's problem was obtained by Doob (1954) by considering a first exit time ; here a similar solution to the so-called Robin's problem will be obtained by considering a last exit time.

Consider the general setting described in § 1. Call a Borel subset B of E "transient" iff for almost every path ω, there is a finite time $t^*(\omega)$ such that $X_t(\omega) \notin B$ for $t > t^*(\omega)$. Define

$$\Delta_B = \{\omega \in \Omega \mid \exists t > 0 : X_t(\omega) \in B\} ;$$

$$\gamma_B(\omega) = \begin{cases} \sup \{t > 0 \mid X_t(\omega) \in B\} & \text{if} \quad \omega \in \Delta_B \ ; \\ 0 \, , \text{ if } \quad \omega \in \Omega - \Delta_B . \end{cases}$$

Then B is transient if and only if $\gamma_B < \infty$ a.s. (almost surely). It is easy to see that γ_B is a random variable, called the *last exit time* from B.

Let us begin by supposing all paths continuous. It then follows that

$$X(\gamma_B) \in \partial B \quad \text{a.s.} \tag{2}$$

From now on we fix B and write γ for γ_B.

It is well-known that a compact set is transient for Brownian motion in R^3 ; thus the setting above includes the Newton-Gauss case described above. It also includes, e.g. the case of a Brownian motion in R^2, terminated after the path leaves an open ball. In this case the state space E should be the open ball with a one-point compactification ∂, and we must assume that $\partial \notin \overline{B}$, the closure of B.

To study the distribution of the last exit position $X(\gamma)$, we put

$$L(x, A) = \underline{\underline{P}}^x\{\gamma > 0 \ ; \ X(\gamma) \in A\} , \quad x \in E \quad , \quad A \in \mathcal{E}. \tag{3}$$

We are going to determine this by calculating

$$\int_{\partial B} L(x, dy) f(y) = \underline{\underline{E}}^x\{\gamma > 0 \ ; f(X(\gamma))\} \tag{4}$$

for every x and every $f \in C_b$ (the class of bounded continuous functions on E), where $\underline{\underline{E}}^x$ denotes the mathematical expectation corresponding to $\underline{\underline{P}}^x$. This is done by a little device as follows. Take any $\varepsilon > 0$ and consider the "mixed approximation" :

$$\frac{1}{\varepsilon}\int_0^\infty \underline{\underline{E}}^x\{f(X_t)\ ;\ t<\gamma\leqslant t+\varepsilon\}\,dt$$
$$=\frac{1}{\varepsilon}\int_0^\infty \underline{\underline{E}}^x\{f(X_t)\,\underline{\underline{P}}^{X_t}[0<\gamma\leqslant\varepsilon]\}\,dt. \quad (5)$$

Setting

$$\psi_\varepsilon(x)=\frac{1}{\varepsilon}\,\underline{\underline{P}}^x\,[0<\gamma\leqslant\varepsilon]$$

we may rewrite (5) as

$$\underline{\underline{E}}^x\left\{\int_0^\infty (f\psi_\varepsilon)\,(X_t)\,dt\right\}=\int_0^\infty\frac{1}{\varepsilon}\int_{\gamma\in(t,t+\varepsilon)} f(X_t)\,d\underline{\underline{P}}^x\,dt$$
$$=\int_{[\gamma>\varepsilon]}\frac{1}{\varepsilon}\int_{\gamma-\varepsilon}^{\gamma} f(X_t)\,dt\,d\underline{\underline{P}}^x+\int_{[0<\gamma\leqslant\varepsilon]}\frac{1}{\varepsilon}\int_0^\gamma f(X_t)\,dt\,d\underline{\underline{P}}^x\ .$$

The last-written integral is bounded by

$$\int_{[0<\gamma\leqslant\varepsilon]}\frac{\gamma}{\varepsilon}\,\|f\|\,d\underline{\underline{P}}^x\leqslant\|f\|\,\underline{\underline{P}}^x\{0<\gamma\leqslant\varepsilon\},$$

which converges to 0 as $\varepsilon\downarrow 0$. On the other hand,

$$\lim_{\varepsilon\downarrow 0}\int_{\gamma-\varepsilon}^{\gamma} f(X_t)\,dt=f(X_\gamma) \quad (7)$$

boundedly, by the continuity of $t\to f(X_t)$. Hence as $\varepsilon\downarrow 0$ the first integral in the last member of (6) converges to $\int_{[\gamma>0]} f(X_\gamma)d\underline{\underline{P}}^x$, which is just the number in (4).

Define the potential U by

$$U\varphi(x)=\underline{\underline{E}}^x\left\{\int_0^\infty\varphi(X_t)\,dt\right\}$$

where φ is any positive measurable function. The result of calculations in (6) is then as follows :

$$\lim_{\varepsilon\downarrow 0}U(f\psi_\varepsilon)\,(x)=\int_{\partial B}L(x\ ,dy)\,f(y). \quad (8)$$

We now make the following assumptions on the potential kernel U. There exists a σ-finite measure ξ such that

i) $U(x, A) = \int_A u(x, y)\, \xi(dy)$ where $u(x, y) \geqslant 0$, for $x \in E$, $A \in \mathcal{E}$;

ii) $y \to u(x, y)^{-1}$ is finite continuous, for $x \in E$;

iii) $u(x, y) = +\infty$ if and only if $x = y$.

The key formula (8) will now be recorded as follows :

$$\forall f \in C_b : \lim_{\epsilon \downarrow 0} \int_E u(x, y)\, \psi_\epsilon(y) f(y)\, \xi(dy) = \int_{\partial B} L(x, dy) f(y) \quad (9)$$

Set

$$M_\epsilon(A) = \int_A \psi_\epsilon(y)\, \xi(dy) \quad , \qquad A \in \mathcal{E}. \tag{10}$$

For any $x \in E$ and $\varphi \in C_K$ (class of continuous functions on E with compact supports), the function

$$y \to \varphi(y)\, u(x, y)^{-1}$$

belongs to C_K by assumption ii) above. Substituting this for f in (9), we have

$$\lim_{\epsilon \downarrow 0} \int_E \varphi(y)\, M_\epsilon(dy) = \int_{\partial B} L(x, dy)\, u(x, y)^{-1} \varphi(y).$$

This being true for every $\varphi \in C_K$, we conclude first that the measures M_ϵ converge vaguely to a measure $\mu (= \mu_B)$ as $\varepsilon \downarrow 0$; and secondly that this vague limit is identified as follows :

$$\forall x \in E : \mu(dy) = L(x, dy)\, u(x, y)^{-1}. \tag{11}$$

It follows from assumption ii) that μ is a σ-finite measure in $\mathcal{E}$. Since $L(x, \cdot)$ is concentrated on ∂B, so is μ. Since $u(x, y) < \infty$ for $x \neq y$ by assumption iii), we have if $x \notin \partial B$:

$$L(x, dy) = u(x, y)\, \mu(dy). \tag{12}$$

This means if $x \notin \partial B$ and $A \subset \partial B$, $A \in \mathcal{E}$, we have

$$L(x, A) = \int_A u(x, y)\, \mu(dy). \tag{13}$$

Now it is clear that for a transient set B we have

$$\{0 \leqslant T_B < \infty\} = \{0 < \gamma_B < \infty\} .$$

Thus if $x \notin \partial B$, the Gauss-Riesz-Frostman formula (1) is just the particular case of (13) for $A = E$. The measure μ_B is called the ***equilibrium measure*** for B, and its total mas $\mu_B(\partial B)$ the ***capacity*** of B, up to a multiplicative constant.

We shall establish (13), and consequently (1), for all $x \in E$. Taking $x = y$ in (1) and using assumption iii), we see that μ is atomless, namely for every $y \in E$:

$$\mu(\{y\}) = 0.$$

Next we have again by iii) and (11), if $x \neq y$:

$$L(x, \{y\}) = u(x, y)\, \mu(\{y\}) = 0 . \tag{14}$$

Finally,

$$L(x, A\backslash\{x\}) = \int_{A\backslash(x)} u(x, y)\, \mu(dy) = \int_A u(x, y)\, \mu(dy),$$

by the usual convention $\infty \cdot 0 = 0$ when the integrand takes the value $+\infty$ at a point while the corresponding mass at the point is 0. Thus (13) will hold true if and only if

$$\forall x \in E : L(x, \{x\}) = 0. \tag{15}$$

A point x in E is called a ***holding point*** iff almost every path starting at x must remain at x for some strictly positive time. We show that if x is nota holding point, namely if for each $\delta > 0$ we have

$$\underline{\underline{P}}^x\{X(t) = x \quad \text{for} \quad t \in [0, \delta]\} = 0, \tag{16}$$

then (15) is true. The following proof requires only that the process be separable. To simplify notation we may then suppose that the dyadic numbers $\left\{\frac{n}{2^m} ; m \geqslant 0, n \geqslant 0\right\}$ are a separability set. For each n we define

$$S_n = \min\left\{\frac{m}{2^n} \mid X\left(\frac{m}{2^n}\right) \neq x\right\}$$

with the convention that $S_n = +\infty$ when the set on the right side is empty. If x is not a holding point, then

$$\underline{\underline{P}}^x\{S_n \downarrow 0\} = 1$$

by separability. We have, therefore, using only ordinary Markov property because S_n is countably-valued :

$$\underline{\underline{P}}^x\{\gamma > S_n \ ; X(\gamma) = x\} \leqslant \underline{\underline{E}}^x\{P^{X(S_n)}[\gamma > 0 \ ; X(\gamma) = x]\} = 0$$

by (14), since $X(S_n) \neq x$. Letting $n \to \infty$ we obtain (15).

On the other hand, if x is a holding point, then

$$U(x\,,\{x\}) = u(x\,,x)\,\xi(\{x\}) > 0\ ,$$

hence $\xi(\{x\}) > 0$ and so $U(x\,,\{x\}) = \infty$ by assumption iii) above. [I owe this observation to Hans Föllmer, which enabled me to deal with a holding point.] This implies that the singleton $\{x\}$ constitutes a "recurrent set" in the sense that starting from x almost every path will hit $\{x\}$ after an arbitrarily large t. A familiar argument then shows that almost no path can lead from x to the transient set B. Thus by definition

$$L(x\,,\{x\}) \leqslant P^x\{T_B < \infty\} = 0$$

and so (15) is also true for a holding point x.

We have therefore established the fundamental result (13) for every x and every A under the hypotheses specified.

Note that the existence of the measure μ_B for the representation given in (1) has been established for every transient set B, without any regularity condition whatever on ∂B. However, the hitting probability on the left side of (1) need not be equal to one for all $x \in B$, as required by the classical definition of equilibrium potential. Herein lies the necessity of a condition like Poincaré's to ensure that every x on ∂B is regular for B, or the exception of a set of points on ∂B which are irregular for B.

Finally, formula (11) gives an explicit solution to Robin's problem of determining the equilibrium measure. Indeed, by a suitable choice of the arbitrary point x there, the probabilistic expression may even yield a deterministic one. A trivial case in point is when B is a ball and x its center.

The method of proof will now be extended to the case where the paths are right continuous with left limits. Here are the necessary changes. Relation (2) is replaced by

$$X(\gamma_B -) \in \overline{B},$$

where $\overline{B}$ is the closure of B. The last exit distribution in (3) is redefined by using the left limit $X(\gamma -)$ instead of $X(\gamma)$. Both $L(x, \cdot)$ and μ are now concentrated on $\overline{B}$ instead of ∂B. If we replace $X(\gamma)$ by $X(\gamma -)$ and ∂B by $\overline{B}$ in the obvious places, all the steps go through as before. The final result is, for each Borel Set $A \subset \overline{B}$, and each $x \in E$:

$$\underline{\underline{P}}^x\{X(\gamma_B -) \in A\} = L(x, A) = \int_A u(x, y)\, \mu_B(dy) \ ; \qquad (17)$$

in particular

$$\underline{\underline{P}}^x\{T_B < \infty\} = L(x, \overline{B}) = \int_{\overline{B}} u(x, y)\, \mu_B(dy).$$

In this form the result covers the M. Riesz-Frostman potentials where $u(x, y) = c/|x - y|^\alpha$, α real > 0, $c =$ constant. As is known, the corresponding Markov process is a stable process whose paths may be assumed to be right continuous and have left limits.

The assumption ii) is expedient for our method and remains to be analysed. Next, we examine some possibilities of relaxing the assumption iii) to illustrate the relationship between the poles of u and the polar-like sets of the process. We shall shun the standard duality assumptions but lead up to them. Let us put

$$N_x = \{y \in E \mid u(x, y) = +\infty\},$$

$$\widetilde{N}_y = \{x \in E \mid u(x, y) = +\infty\}.$$

Thus our previous assumption iii) amounts to the simplest of its kind :

$$N_x = \widetilde{N}_x = \{x\}.$$

We will not assume this, but confine ourselves to a Hunt process below in order to avail ourselves of standard arguments and results of the theory.

0) If for every x, $N_x = \emptyset$, then (17) is true by (11).

I) Suppose that for each $y \in \overline{B}$, $\widetilde{N}_y \neq \emptyset$. Then in order that (17) be true as asserted, it is necessary that μ_B is atomless ; and it is necessary and sufficient that

$$\forall x \in \overline{B} : L(x , N_x \cap \overline{B}) = 0. \tag{18}$$

Proof. – Take $y \in \overline{B}$, $x \in \widetilde{N}_y$, $A = \{y\}$ in (17) :

$$L(x , \{y\}) = u(x , y)\, \mu_B(\{y\}) ,$$

hence $\mu_B(\{y\}) = 0$. On the other hand, if (18) holds, then for every $A \subset \overline{B}$:

$$L(x , A) = L(x , A \backslash (N_x \cap \overline{B})) = \int_{A \backslash (N_x \cap \overline{B})} u(x , y)\, \mu(dy)$$

$$= \int_A u(x , y)\, \mu(dy),$$

where the second equation is true by (11), the third because $\mu(N_x) = 0$ from

$$1 \geqslant L(x , N_x) = \int_{N_x} (+\infty)\, \mu(dy) .$$

The necessity of (18) is shown in the same way.

II) Suppose for each $y \in \overline{B}$, $\widetilde{N}_y \neq \emptyset$ and $E - \widetilde{N}_y$ is finely dense (in particular if $\widetilde{N}_y$ is of null U-potential), then for each x, $L(x , \cdot)$ is atomless. If in addition, the set $N_x \cap \overline{B}$ is countable for each $x \in \overline{B}$, then (17) is true as asserted.

Proof. – It follows as before that if $x \notin \widetilde{N}_y$, then $L(x , \{y\}) = 0$. This is then true for every x by an argument similar to that given after (16), based on the fact that x is regular for $E - \widetilde{N}_y$. When $N_x \cap \overline{B}$ is countable, (18) follows.

III) Suppose that for each x , $y \to u(x , y)$ is an excessive function, and $E - N_x$ is finely dense. Then (17) is true as asserted.

Proof. – For each $y \notin N_x$, the process $\{u(x , X_t)\ ,\ t \geqslant 0]$ is a supermartingale under $\underline{\underline{P}}^y$. Hence by a well known result of Doob's :

$$\underline{\underline{P}}^y\{\exists t > 0 : X_t \in N_x\} = 0 ;$$

this is then also true if we replace X_t by X_{t-}, by a result of Hunt's. Furthermore, the result is true for every y, since $E - N_x$ is finely dense. It follows that $L(x, N_x) = \underline{\underline{P}}^x\{\gamma > 0 ; X(\gamma -) \in N_x\} = 0$ for every x.

Of course the assumption $y \to u(x, y)$ is "wrong", since generally it is $x \to u(x, y)$ that should be excessive, whereas $y \to u(x, y)$ should be co-excessive. However, if $u(x, y)$ is symmetric in (x, y) this makes sense. Let us also remark that if $y \to u(x, y)$ is locally integrable with respect to the *reference* measure ξ, then $\xi(N_x) = 0$. Consequently, N_x is of null potential and so $E - N_x$ is indeed finely dense as assumed.

REFERENCES

[1] M. BRELOT, Les étapes et les aspects multiples de la théorie du potentiel, *L'Enseignement mathématique*, **58** (1972), 1-36.

[2] K. ITO and H.P. MCKEAN, *Diffusion Processes and Their Sample Paths*, Springer-Verlag, 1965.

[3] H.P. MCKEAN, A probabilistic interpretation of equilibrium charge distribution, *J. Math. Kyoto Univ.*, 4 (1965), 617-623.

Manuscrit reçu le 26 septembre 1972

Accepté par M. Brelot

K.L. CHUNG

Department of Mathematics

Stanford University

Stanford, California 94305 (USA)

Z. Wahrscheinlichkeitstheorie verw. Gebiete 29, 253–256 (1974)

Meyer's Theorem on Predictability

Kai Lai Chung* and John B. Walsh

In a Hunt process the sample path is almost surely continuous at a predictable time, because of quasi left continuity. Meyer proved the striking converse that an optional time is predictable if the sample path is almost surely continuous thereat. His proof [1], [2; 111–116] may be divided into three steps: (i) every bounded martingale is continuous where the path is; (ii) given a totally inaccessible time one can construct a bounded martingale which has a unique jump thereat; (iii) for a Hunt process an accessible time is predictable. The purpose of this note is to give a simpler proof of Meyer's theorem (Proposition 4 here). First we prove a general lemma on the "left substitutability" along the path, after which (i) can be proved with minimal calculations. Originally we have considered a potential for the h in (1) below, and it was Meyer who in turn pointed out the ultimate simple form of the result. Next, under the hypothesis that the path is continuous at T, we actually construct a sequence of optional times announcing T. It is a tantalizing thought that we need (i) only for the one martingale $E\{T|\mathfrak{F}_t\}$, but we do not know how to show the continuity of this martingale directly. Perhaps an intuitively obvious proof will be found someday.

We are limiting ourselves to Hunt processes for clarity, but the result holds with slight modifications for standard processes (see [6]), and our proof can be easily extended to encompass this case.

We consider a Hunt process $X=(\Omega, \mathfrak{F}, \mathfrak{F}_t, X_t, \theta_t, P^x)$ taking values in a locally compact separable metric space E, and we adopt the standard notations and conventions as used in [2] and [3], except for the following abuse of notation: the standard definitions of predictability, inaccessibility, etc., are relative to one fixed probability law, which in this case is a law P^μ. We shall omit the initial distribution μ and write, for instance, $E\{X\}$ and $\mathfrak{F}_t$ instead of $E^\mu\{X\}$ and $\mathfrak{F}_t^\mu$. The scrupulous reader has the option of inserting the numerous superscripts and the accompanying explanatory phrases. It is perhaps worthwhile to emphasize that the σ-fields $\mathfrak{F}_t$ are those generated by the pocess and completed in the usual manner with respect to P $(=P^\mu)$, and that they are right continuous in t. A martingale relative to $(\mathfrak{F}_t)$ has a version which is almost surely right continuous and has left limits everywhere as a function of t. A useful fact is that if M_t is a well-measurable martingale satisfying the stopping rule (i.e. $M_T=E\{M_{t_0}|\mathfrak{F}_T^\mu\}$ whenever $T\leqq t_0$ is optional), then M is already P-almost surely right continuous.

We write "$s\uparrow\uparrow t$" to mean "$s<t$ and s increases to t". Thus, T is a predictable time iff there exists a sequence of optional times $\{T_n\}$ such that $T_n\uparrow\uparrow T$ on $\{T>0\}$; such a sequence is said to *announce* T.

* Research supported in part by NSF grant GP-41710X.

Proposition 1. *Let h be a bounded Borel function on E and let $0<t_1<\infty$. Then we have a.s.*

$$\forall\, t\in(0,t_1]:\ \lim_{s\uparrow\uparrow t} P_{t_1-s}h(X_s)=P_{t_1-t}h(X_{t-}). \tag{1}$$

Proof. If T is optional and $T\leqq t_1$, then we have

$$E\{h(X_{t_1})|\mathfrak{F}_T\}=P_{t_1-T}h(X_T). \tag{2}$$

Here we have used a form of the strong Markov property (see [2; 81]), which is relatively unused and turns out to be essential for our purpose.

Since X is well measurable and $(s,x)\to P_{t_1-s}h(x)$ is Borelian, it is easily verified that the process

$$P_{t_1-s}h(X_s), \quad 0\leqq s\leqq t_1, \tag{3}$$

is well measurable. Eq. (2) shows that (3) is a martingale which satisfies the stopping rule, hence it is right continuous and has left limits. In particular, the limit in the left member of (1) exists and defines a left-continuous, hence predictable, process. The right member of (1) is also predictable, being a Borel function of the left continuous process (X_{t-}). Therefore the two members of (1) will be indistinguishable if we show that they are equal a.s. at an arbitrary predictable time; see [4; T13, 73]. Let $T\leqq t_1$ be such a time, and let $\{T_n\}$ announce T. Applying (2) and the martingale convergence theorem, we have

$$\begin{aligned}\lim_{n\to\infty} P_{t_1-T_n}h(X_{T_n}) &= \lim_{n\to\infty} E\{h(X_{t_1})|\mathfrak{F}_{T_n}\}\\ &= E\{h(X_{t_1})|\mathfrak{F}_{T-}\}\end{aligned} \tag{4}$$

since $\bigvee_{n=1}^{\infty}\mathfrak{F}_{T_n}=\mathfrak{F}_{T-}$. Now by quasi left continuity $X_T=\lim_{n\to\infty}X_{T_n}=X_{T-}$, which belongs to $\mathfrak{F}_{T-}$. Since $\mathfrak{F}_{T-}\subset\mathfrak{F}_T$, this implies that (2) remains true when the field $\mathfrak{F}_T$ is replaced by $\mathfrak{F}_{T-}$.[1] Thus the limit in (4) is a.s. equal to the right member in (2) where we may read X_{T-} for X_T. In other words (1) is true when t is replaced by T, as was to be shown. Proposition 1 is proved.

Remark. Proposition 1 is equivalent to the second part of the following assertion, of which the first part is proved in a similar way. We have a.s.: $t\to P_{t_1-t}h(X_t)$ is right continuous in $[0,t_1)$, and $t\to P_{t_1-t}h(X_{t-})$ is left continuous in $(0,t_1]$. Compare with the analogues for excessive functions and potentials in [2].

Proposition 2. *Let $\{M_t,\mathfrak{F}_t: t\geqq 0\}$ be a right continuous martingale. Then for a.e. ω, $t\to M_t(\omega)$ is continuous wherever $t\to X_t(\omega)$ is.*

Proof. It is sufficient to prove this for $t\in[0,t_1]$ for each $t_1>0$. Since $M_t=E\{M_{t_1}|\mathfrak{F}_t\}$ for $t\in[0,t_1]$, it is sufficient to consider martingales of the form $E\{Y|\mathfrak{F}_t\}$ where $Y\in L^1(\Omega,\mathfrak{F},P)$. A set $D\subset L^1$ is said to be *fundamental* iff finite linear combinations of the form $\sum_{j=1}^{l} c_j Y_j$, where $Y_j\in D$, are dense in L^1. It is sufficient to consider Y in such a set, since if $Y_n\to Y$ in L^1, then $\sup_t E\{|Y_n-Y|\,|\mathfrak{F}_t\}\to 0$ in probability by the maximal inequality for martingales. All this is as in Meyer's proof but now we

[1] In fact, it is known that $\mathfrak{F}_T=\mathfrak{F}_{T-}$ here.

use the simple fundamental set given by random variables of the form:

$$Y=\prod_{i=1}^{n} f_i(X_{t_i}) \tag{5}$$

where each f_i is bounded Borelian, and $0=t_0<t_1<\cdots<t_n<t_{n+1}=\infty$. For this Y we have, if $t\in[t_k,t_{k+1}]$, $1\leqq k\leqq n$:

$$E\{Y|\mathfrak{F}_t\}=\prod_{i=1}^{k} f_i(X_{t_i})\,E\left\{\prod_{i=k+1}^{n} f_i(X_{t_i})\middle|\mathfrak{F}_t\right\},$$

where the last-written expectation is to be omitted for $k=n$. It follows that it is sufficient to consider, for the Y in (5),

$$M_t=E\{Y|\mathfrak{F}_t\} \quad \text{for } t\in[0,t_1]. \tag{6}$$

Put

$$h=f_1\prod_{i=2}^{n} P_{t_i-t_{i-1}}f_i.$$

We have by the Markov property:

$$M_t=P_{t_1-t}\,h(X_t).$$

It follows from Proposition 1 that

$$M_{t-}=P_{t_1-t}\,h(X_{t-}).$$

If X is continuous at t, we may replace X_{t-} above by X_t and the result is then $M_{t-}=M_t$. Hence M is continuous at t and the proposition is proved.

Define the supermartingale which is right continuous and has left limits:

$$Z_t=E\{T-(T\wedge t)|\mathfrak{F}_t\}=E\{T|\mathfrak{F}_t\}-(T\wedge t), \quad 0\leqq t<\infty. \tag{7}$$

Proposition 3. *We have a.s.*

(i) $\quad Z_{t-}>0$ *and* $Z_t>0$ *on* $\{0<t<T\}$;

(ii) $\quad Z_t=0$ *on* $\{t\geqq T\}$.

Proof. For a fixed t let

$$\Lambda=\{t<T;\ Z_t\leqq 0\};$$

then $\Lambda\in\mathfrak{F}_t$ and so

$$\int_\Lambda Z_t\,dP=\int_\Lambda (T-t)\,dP. \tag{8}$$

The integrand on the left is negative while that on the right is strictly positive. Hence $P(\Lambda)=0$, and consequently we have a.s. $Z_r>0$ for all rational $r\in(0,T)$. Now Z is a positive supermartingale; if for some $\tau>0$ either $Z_{\tau-}$ or Z_τ is zero, then Z_t will be zero for all $t\geqq\tau$ (see Meyer [5; T15, 134]). This is impossible by what has just been shown. Therefore (i) is true. To prove (ii), we note first that it is true for all rational values of t by considering $\Lambda=\{t>T;\ Z_t>0\}$ in (8), and then conclude by right continuity of Z.

Proposition 4. *An optional time T is predictable if and only if X is continuous at T a.s. on $\{T<\infty\}$; T is totally inaccessible if and only if X has a jump at T a.s. on $\{T<\infty\}$.*

Proof. We may suppose $E\{T\}<\infty$ by replacing T with $T\wedge N$ and then letting $N\to\infty$.

Suppose that X is continuous at T. Define

$$T_n=\inf\left\{t>0\colon Z_t\leqq\frac{1}{n}\right\}$$

where Z is as in (7). Clearly T_n is optional and $T_n\leqq T_{n+1}$ for all n. Now $t\to Z_t$ is a.s. continuous at T because $E\{T|\mathfrak{F}_t\}$ is by Proposition 2, and $Z_T=0$ by Proposition 3. This clearly implies that $T_n<T$ for all n. Put $T_\infty=\lim_{n\to\infty} T_n$, then $T_\infty\leqq T$. Since $Z(T_n)\leqq 1/n$ by right continuity, and $Z(T_\infty)=\lim_{n\to\infty} Z(T_n)=0$ by the existence of left limit, we must have $T_\infty=T$ by Proposition 3. We have therefore proved that T is predictable, in fact $\{T_n\}$ announces T.[2] Conversely if T is predictable, then X is a.s. continuous at T on $\{T<\infty\}$, as mentioned in the opening sentence.

Suppose that X has a jump at T on $\{T<\infty\}$. Let S be any predictable time announced by $\{S_n\}$, and $\Lambda=\{0<S=T<\infty\}$. Then on Λ, $S_n\uparrow\uparrow T$ and so by quasi left continuity we have $X_{T-}=X_T$. The assumption that X has a jump at T on Λ therefore entails $P(\Lambda)=0$. This being true for every predictable S, T is by definition totally inaccessible. Conversely, suppose that T is totally inaccessible. Let $\Lambda=\{X_{T-}=X_T;\ T<\infty\}$, then $\Lambda\in\mathfrak{F}_T$ and T_Λ (defined to be T on Λ; and ∞ on $\Omega-\Lambda$) is optional. Obviously X is continuous on $\{T_\Lambda<\infty\}$, and so by our main result T_Λ is predictable. Since $T=T_\Lambda<\infty$ on Λ, this entails $P(\Lambda)=0$, which means X has a jump a.s. at T on $\{T<\infty\}$. Proposition 4 is completely proved.

References

1. Meyer, P. A.: Decomposition of supermartingales: the uniqueness theorem. Illinois J. Math. 7, 1-17 (1963)
2. Meyer, P. A.: Processus de Markov, Lecture Notes in Mathematics, Vol. 26. Berlin-Heidelberg-New York: Springer 1967
3. Blumenthal, R. M., Getoor, R.K.: Markov processes and potential theory. New York: Academic Press 1968
4. Dellacherie, C.: Capacités et processus stochastiques. Berlin: Springer 1972
5. Meyer, P. A.: Probabilités et potential. Paris: Hermann 1966
6. Getoor, R. K.: Regularity of excessive functions II. Ann. Math. Statist. 42, 2057-2063 (1971)

John B. Walsh
University of British Columbia
Department of Mathematics
Vancouver
British Columbia/Canada

Kai Lai Chung
Department of Mathematics
Stanford University
Stanford, California 94305
USA

(Received April 19, 1974)

[2] An interesting consequence of this fact: assume T is optional and predictable with respect to any family $(\mathfrak{F}_t)=(\mathfrak{F}_t^\mu)$ and law $P=P^\mu$. Then it is known that there is one single process which is a right continuous version of the martingale $E\{T|\mathfrak{F}_t\}$ for any law P. Using this process in the definition (7), we get one single family (T_n) which announces T for all these laws simultaneously.

APPROXIMATION OF A CONTINUOUS PARAMETER MARKOV CHAIN BY DISCRETE SKELETONS

KAI LAI CHUNG

The approximation of a given continuous parameter Markov chain by its discrete skeletons raises many interesting questions. Some beginnings of this study were given in §II.13 of [1], but little progress seems to have been made since. It is possible that the apparent impenetrability of the "embedding problem" (when is a discrete parameter chain a skeleton?) may have unwarrantedly hampered reasonable and useful investigations. The present note deals with a relatively easy aspect of the approximation theory: the global faithfulness of the skeletal approach as reflected in the moments of the entrance time distributions. Even so our method leans heavily on recurrence. When this hypothesis is dropped these distributions may be improper, but moments can still be defined by excluding the mass at infinity as done in §I.11 of [1] in the discrete parameter case. Solidarity results such as Theorem 1 there can be generalized to the continuous parameter case without the intervention of skeletons (see the remarks concerning the last part of the theorem below); but approximation seems an open problem.

The basic elements of continuous parameter Markov chains as given in [1] are assumed here, but the main arguments should be accessible to readers with a casual acquaintance with the subject. Notations in [1] have been modified to accord with the current usage.

Let $\{X_t, t \geqslant 0\}$ be a continuous parameter Markov chain in the right lower semi-continuous version (see [1; §II.7]). For $h > 0$ let $\{X_n^{(h)}, n \geqslant 0\}$ be its h-skeleton, i.e., $X_n^{(h)} = X_{nh}$; this is then a discrete parameter Markov chain. The state space of the latter may be taken to be the minimal state space of the continuous parameter chain. Let j be an arbitrary state, and put

$$T_j(\omega) = \inf\{t > 0 \mid X_t(\omega) = j\},$$
$$T_j^{(h)}(\omega) = \inf\{n \geqslant 1 \mid X_n^{(h)}(\omega) = j\}\, h.$$

These are the first entrance times into j for $\{X_t\}$ and $\{X_n^{(h)}\}$ respectively. It is proved in [1; §II.13] that for each given sequence $\{h_\nu\} \downarrow 0$, we have

$$\lim_{\nu\to\infty} T_j^{(h_\nu)} = T_j, \quad P_i\text{-a.s.} \tag{1}$$

for every initial state i. But if j is instantaneous, then the set of real numbers h for which $T_j^{(h)}$ is finite a.s. constitutes only a set of the first category, so that (1) becomes decidedly false if the sequential convergence $h_\nu \downarrow 0$ is replaced by plain convergence as $h \downarrow 0$. This is the kind of ill behaviour exhibited by an instantaneous state which makes the following result noteworthy.

Received 5 November, 1973; revised 11 January, 1974.

[Bull. London Math. Soc., 7 (1975), 71–76]

THEOREM. *Let the state space form a recurrent class. Suppose that for two distinct states i and j, and some real number $r > 0$, we have*

$$\mathbf{E}_i\{T_j^r\} < \infty, \quad \mathbf{E}_j\{T_i^r\} < \infty. \tag{2}$$

Then for every $h > 0$, $\mathbf{E}_i\{(T_j^{(h)})^r\} < \infty$ and

$$\lim_{h \to 0} \mathbf{E}_i\{|T_j^{(h)} - T_j|^r\} = 0. \tag{3}$$

Moreover, (2) is then true for any two distinct states i and j; and (3) is true for any j and any initial state i. Finally, if there is a stable state i in the recurrent class, then the conditions in (2) may be replaced by a single one:

$$\mathbf{E}_i\{R_i^r\} < \infty; \tag{4}$$

where R_i is defined as follows:

$$Q_i = \inf\{t > 0 \mid X_t \neq i\},$$

$$R_i = \inf\{t > Q_i \mid X_t = i\}.$$

The essential part of the theorem is of course (3). Once this is proved, then we know that for some h, $\mathbf{E}_i\{(T_j^{(h)})^r\} < \infty$ and $\mathbf{E}_j\{(T_i^{(h)})^r\} < \infty$, the latter by symmetry. Applying known results for a discrete parameter chain ([1; Theorem II.11.1], we see that these relations are also true for any two distinct states i and j. Since it is trivial by definition that for any j and any $h > 0$:

$$T_j \leqslant T_j^{(h)},$$

the conditions in (2) are then true for any two distinct states and so the proven conclusion (3) will apply to them as well. Finally if (4) is assumed then the arguments spelled out in the discrete parameter case on p. 63 of [1] can be used (without change except from discrete distributions to general ones) to deduce (2) for any j distinct from i, so that we are back to the original assumptions.

We proceed to the proof of (3). Using the taboo probability function ${}_ip_{jj}$, we begin by choosing a sequence of positive numbers $\{\delta_v\} \downarrow 0$ such that

$$h < \delta_v \Rightarrow {}_ip_{jj}(h) > 1 - (1/2^v). \tag{5}$$

This is possible since $\lim_{t \to 0} {}_ip_{jj}(t) = 1$ (see Theorem 2 in §II.11 of [1]). It is sufficient to prove (3) for a sequence $\{h_v\}$ satisfying $h_{v+1} < h_v$ and $h_v < \delta_v$ for all $v \geqslant 1$. For then given any sequence converging to zero, there is a subsequence along which (3) holds. It then follows from the usual compactness argument that (3) holds for any sequence converging to zero. Hence it holds as asserted by the nature of convergence for real numbers. This is to be contrasted with the remarks following (1).

Suppose that $X_0 = i$. Let S_1 be the first entrance time into j, S_1' be the first entrance time into i after S_1; in general for $n \geqslant 2$:

$$S_n = \text{first entrance time into } j \text{ after } S'_{n-1},$$

$$S_n' = \text{first entrance time into } i \text{ after } S_n.$$

Put also for $n \geqslant 1$:

$$T_n = S_{n+1} - S_n.$$

The random variables $\{T_n\}$ are independent and identically distributed. It follows from the hypothesis (2) that

$$\mathbf{E}\{T_1^r\} < \infty. \tag{6}$$

Our next step is to select those T_n's which are large enough that the random interval (T_n, T_{n+1}) may contain certain phenomena to be specified below. Such a selection is possible owing to the following lemma.

LEMMA 1. *For any constant* $t_0 > 0$ we have

$$0 < P\{T_1 > t_0\} < 1. \tag{7}$$

Proof. Let F_{ij} denote the first entrance time distribution from i to j. Then T_1 has the distribution $F_{ij} {}^* F_{ji}$. It is known that for any $i \neq j$, F_{ij} has a continuous density f_{ij} which is strictly positive in $(0, \infty)$; see Theorem II.12.5 and Theorem II.1.5 of [1], the latter trivially extended to a substochastic transition semigroup. The lemma follows at once from this. Since these results are rather deep, here is an easy proof of the first inequality in (7). Using the obvious probabilistic meaning, we have

$$1 - F_{ij}(s+t) \geqslant {}_jp_{ii}(s)[1 - F_{ij}(t)].$$

Since

$${}_jp_{ii}(s) \geqslant p_{ii}(s) - \sup_{0 \leqslant t \leqslant s} p_{ij}(t) > 0$$

for sufficient small s, the preceding inequality implies that $F_{ij}(s+t) = 1$ would force $F_{ij}(t) = 1$, which yields a quick contradiction. The second inequality in (7) follows from $p_{ij} \leqslant F_{ij}$ and the cited Theorem II.1.5, without Theorem II.12.5; but it is not really needed below.

It follows from Lemma 1 that $P\{T_1 > 3\} = \alpha > 0$. Put

$$N = \inf\{n \geqslant 1 \mid T_n > 3\}. \tag{8}$$

Since the T_n's are independent and identically distributed, N has the geometric distribution

$$P\{N = n\} = (1-\alpha)^{n-1}\alpha. \tag{9}$$

We know from Lemma 1 that $\alpha < 1$, but if α were equal to one the relation below would be trivial:

$$\mathbf{E}\{(T_1 + \ldots + T_N)^r\} < \infty. \tag{10}$$

This can be proved as an extension of Wald's equation under the assumptions that $\mathbf{E}(T_1^r) < \infty$ and $\mathbf{E}(N^r) < \infty$. I was unable to find such a result in the literature but Louis Gordon supplied me with a proof based on the inequalities of Marcinkiewicz and Zygmund. A direct proof was also obtained for integral r by David Siegmund. The interested reader is invited to communicate with them for this nice result, but we will make do with the more crude lemma below which is trivial but applicable also later on; it was used in the proof of Theorem I.14.4 of [1].

LEMMA 2. *Suppose that the random variables $\{X_n, n \geqslant 1\}$ have finite moments of order $r > 0$; N $(\geqslant 1)$ is an integer-valued random variable, and there is a constant $C < \infty$ such that for all $n \geqslant 1$ we have*

$$\max_{1 \leqslant k \leqslant n} \mathbf{E}\{|X_k|^r \mid N = n\} \leqslant C, \tag{11}$$

(*where an undefined conditional expectation is to be omitted*).
Then

$$\mathbf{E}\left\{\left|\sum_{k=1}^{N} X_k\right|^r\right\} \leqslant C\mathbf{E}(N^r).$$

To apply this lemma to the situation in (10), we note that

$$\{N = n\} = \bigcap_{k=1}^{n-1} \{T_k \leqslant 3\} \cap \{T_n > 3\};$$

and since the T_n's are independent in this case, we have

$$\mathbf{E}\{T_k^r \mid N = n\} = \mathbf{E}\{T_k^r \mid T_k \leqslant 3\}, \qquad 1 \leqslant k \leqslant n-1;$$

$$\mathbf{E}\{T_n^r \mid N = n\} = \mathbf{E}\{T_n^r \mid T_n > 3\}.$$

Since for each k,

$$P\{T_k \leqslant 3\} = 1-\alpha, \qquad P\{T_k > 3\} = \alpha,$$

and the T_n's have the same distribution as T_1, the inequality (11) is trivially satisfied if we take

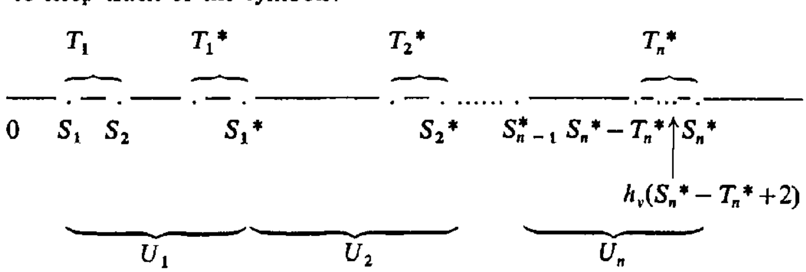
$$C = \max\left(\frac{1}{\alpha}, \frac{1}{1-\alpha}\right)\mathbf{E}\{T_1^r\} < \infty.$$

Having proved (10) by Lemma 2, we put $N_1 = N$ and for $k \geqslant 1$:

$$N_{k+1} = \inf\{n > N_k \mid T_n > 3\},$$

$$T_k^* = T_{N_k}, \qquad S_k^* = S_1 + \sum_{n=1}^{N_k} T_n. \tag{12}$$

Thus, the sequence $\{T_k^*, k \geqslant 1\}$ is that of the selected T_n's larger than 3, and S_k^* is the end-point of the k-th selected interval whose length is T_k^*. A diagram should help to keep track of the symbols:



We shall also use the following notation. For each $h > 0$ define $h(x)$ for all $x > 0$:

$$h(x) = [x/h]\,h+h,$$

where $[.]$ is the greatest integer function, so that $h(x)$ is the smallest integral multiple of h that is strictly larger than x. By definition, the chain is in the state j at the epoch $S_n^* - T_n^*$. We want to impose the event that its h_ν-skeleton, for all $\nu \geqslant 1$, will enter j in the time interval $(S_n^* - T_n^*, S_n^*)$, but not too near the left end-point. The last proviso is made to ensure that the conditional probability of the said event be bounded away from one as well as from zero, so that Lemma 2 may be applied. Of course this event may or may not occur for each n, but it eventually will occur owing to recurrence. Here is the event in question:

$$E_n = \{X(h_\nu(S_n^* - T_n^* + 2)) = j \quad \text{for all} \quad \nu \geqslant 1\}. \tag{13}$$

Note that E_n implies $X(S_n^* - T_n^* + 2) = j$ by right lower semi-continuity. We may suppose that $h_1 = 1$. Then E_n is determined by $X(t)$ with $t \in (S_n^* - T_n^*, S_n^*)$ since $T_n^* > 3$. These intervals are disjoint and their end points are optional times at which X takes the value j. Hence the strong Markov property implies that the events $\{E_n, n \geqslant 1\}$ are independent. We show that there exist two constants α_1 and α_2 such that for all $n \geqslant 1$:

$$0 < \alpha_1 \leqslant P(E_n) \leqslant \alpha_2 < 1. \tag{14}$$

Under P_j let θ be the first entrance time into i, and θ' be the first entrance time after θ into j. Remember that $X(S_n^* - T_n^*) = j$, $T_n^* > 3$, and $X(2) = j$ implies $\theta > 2$. We have

$$\begin{aligned} P(E_n) &\leqslant P_j\{X(2) = j \mid \theta' > 3\} \\ &= \frac{P_j\{X(2) = j; \theta > 3\}}{P_j\{\theta' > 3\}} \leqslant \frac{P_j\{\theta > 2, \theta' > 3\}}{P_j\{\theta' > 3\}} = \alpha_2. \end{aligned}$$

Since θ and $\theta' - \theta$ are independent random variables having respectively F_{ji} and F_{ij} as distributions, it follows from the proof of Lemma 1 that $\alpha_2 < 1$.

Next, let $\{d_\nu(\omega), \nu \geqslant 1\}$ be the non-decreasing rearrangement of distinct members of the sequence $h_\nu(S_n^* - T_n^* + 2) - (S_n^* - T_n^*)$. Then

$$2 < d_\nu(\omega) \leqslant 2 + h_\nu \leqslant 3 \quad \text{for} \quad \nu \geqslant 1, \quad d_\nu(\omega) \downarrow 2,$$

and $d_\nu(\omega) - d_{\nu-1}(\omega) \leqslant h_\nu$. Since d_ν is a function of $S_n^* - T_n^*$, we may apply the strong Markov property in the generality discussed on p. 179 of [1]. We have

$$\begin{aligned} P(E_n) &= P_j\{X(2) = j;\ X(d_\nu) = j, \quad \text{for all } \nu \geqslant 1 \mid \theta' > 3\} \\ &\geqslant P_j\{X(2) = j;\ X(d_\nu) = j, \quad \text{for all } \nu \geqslant 1;\ \theta' > 3\} \\ &\geqslant {}_i p_{jj}(2)\,\mathrm{E}_j\left\{\prod_{\nu=1}^{\infty} {}_i p_{jj}(d_{\nu+1} - d_\nu)\right\} P_j\{\theta' > 3\} = \alpha_1. \end{aligned}$$

Since ${}_ip_{jj}$ is a strictly positive function, $\alpha_1 > 0$ by (5) and Lemma 1. We have proved (14).

Going back to (12), putting $N_0 = 0$, $S_0{}^* = S_1$; and for $k \geqslant 1$:

$$U_k = \sum_{n=N_{k-1}+1}^{N_k} T_n = S_k{}^* - S_{k-1}^*,$$

we have

$$S_n{}^* = S_1 + \sum_{k=1}^{n} U_k.$$

The U_k's are independent and identically distributed, and $\mathbf{E}\{U_1{}^r\} < \infty$ by (10). Now define M to be the smallest value of n for which E_n occurs; thus for each $n \geqslant 1$:

$$\{M = n\} = E_1{}^c \cap \ldots \cap E_{n-1}^c \cap E_n;$$

and by (14)

$$P\{M = n\} \leqslant (1-\alpha_1)^{n-1}\alpha_2$$

so that $\mathbf{E}(M^r) < \infty$. Furthermore we have

$$\mathbf{E}\{U_k{}^r \mid M = n\} = \mathbf{E}\{U_k{}^r \mid E_k{}^c\} \leqslant \frac{1}{1-\alpha_2}\mathbf{E}\{U_k{}^r\}, \qquad 1 \leqslant k \leqslant n-1;$$

$$\mathbf{E}\{U_n{}^r \mid M = n\} = \mathbf{E}\{U_n{}^r \mid E_n\} \leqslant \frac{1}{\alpha_1}\mathbf{E}\{U_n{}^r\}.$$

It follows that $\mathbf{E}\{(S_M{}^*)^r\} < \infty$ by $\mathbf{E}\{S_1{}^r\} < \infty$ and another application of Lemma 2. From the definition of M, we have

$$X(h_\nu(S_M{}^* - T_M{}^* + 2)) = j,$$

and consequently

$$T_j^{(h_\nu)} \leqslant S_M{}^* - T_M{}^* + 3 \leqslant S_M{}^*, \quad \text{for all } \nu \geqslant 1. \tag{15}$$

This means the sequence of random variables $T_j^{(h_\nu)}$ are dominated by $S_M{}^*$ which is in L^r, hence for $h_1 = 1$:

$$\mathbf{E}\{T_j^{(1)}\} < \infty; \tag{16}$$

and since $T_j^{(h_\nu)}$ converges to T_j a.s. as $\nu \to \infty$ by (1), we conclude that the convergence also takes place in L^r, as was to be proved. Finally, consider the time-dilated process $Y(t) = X(ht)$ for an arbitrary $h > 0$. The hypotheseis in (2) are then inherited by the Y process because a change of time unit does not affect the finiteness of these moments. Hence the analogue of (16) for the Y process is also true, which is just (16) itself with the parameter 1 replaced by h. The theorem is completely proved.

Reference

1. Kai Lai Chung, *Markov chains with stationary transition probabilities*, Second Edition (Springer-Verlag, Berlin, 1967).

Maths. Department,
Stanford University,
Stanford, California 94305, U.S.A.

Ann. Inst. Fourier, Grenoble
25, 3 et 4 (1975), 131-138.

REMARKS ON EQUILIBRIUM POTENTIAL AND ENERGY

by Kai Lai CHUNG (*)

Dédié à Monsieur M. Brelot à l'occasion de son 70e anniversaire.

1. To M. Brelot we owe various basic principles and methods of general potential theory (see e.g. [1]). The relationship between them constitutes an important part of the development. In a recent paper [2] I established an equilibrium principle for a broad class of Markov processes by a simple probabilistic method which may be succinctly described as that of « last exit ». In contrast, the probabilistic method of solving the Dirichlet problem, due largely to Doob, may be described as that of the « first exit ». Now the classical method of solving the equilibrium problem, introduced by Gauss and perfected by Frostman, accrues from the minimization of a quadratic functional involving the « energy ». It is natural to ask whether the method of last exit has anything to do with that of energy. Indeed, the first question that arises is whether the equilibrium measure obtained in [2] does minimize some kind of energy. The hypotheses made there are free from the usual duality assumptions and certainly do not require the symmetry of the (potential density) kernel, on which the classical method of energy relies heavily. Although the concept of energy has been extended to the nonsymmetric case, its utilization in a general probabilistic context appears to be still a distant goal.

(*) This work was supported in part by NSF grant GP 41710 at Stanford University.

In this note we shall show that for a symmetric kernel, the equilibrium measure obtained by the last exit method does in fact minimize the energy. Even this little step requires much strengthening of the hypotheses in [2] to be specified below.

2. First we derive a consequence of the results in [2]. Recall that we have a temporally homogeneous Markov process $\{X_t, t \geq 0\}$ taking values in a locally compact topological space E with countable base, and its topological Borel field $\mathcal{E}$. It will be assumed that all paths are continuous; see the last section for remarks concerning the more general case covered in [2]. The potential density u (with respect to some reference measure) satisfies the following conditions:

(a) for each $x \in E$, $u(x, y)^{-1}$ is finite continuous in y;

(b) $u(x, y) = \infty$ if and only if $x = y$.

We will not be concerned with the generalizations of condition (b) given toward the end of [2]. Define for $B \in \mathcal{E}$:

$$\begin{aligned} \text{(1)} \qquad T_B &= \inf\{t > 0 : X_t \in B\}, \\ \gamma_B &= \sup\{t > 0 : X_t \in B\}, \end{aligned}$$

with the convention that $\inf \emptyset = \infty$, $\sup \emptyset = 0$ when $\emptyset$ is the empty set. Then the principal result of [2] is as follows. For each transient set B there exists a σ-finite measure μ_B with support in the boundary ∂B of B, such that for every $x \in E$ we have

$$\text{(2)} \qquad P^x\{T_B < \infty\} = \int_{\partial B} u(x, y)\mu_B\,(dy).$$

Moreover, μ_B is determined in the following way. For $A \in \mathcal{E}$ let

$$\text{(3)} \qquad L_B(x, A) = P^x\{\gamma_B > 0;\ X(\gamma_B) \in A\};$$

then $L_B(x, .)$ has support in ∂B, and

$$\text{(4)} \qquad \mu_B(A) = \int_A u(x, y)^{-1} L_B(x, dy)$$

for any $x \in E$. Thus the right member of (4) does not depend on x.

From now on we assume that each compact is transient; more generally we may consider only transient compacts below. For such a set K we put

$$\text{(5)} \qquad D(K) = \mu_K(E)^{-1}.$$

Lemma. — *Let* $\{K_n,\ n \geq 1\}$ *be compacts such that* $K_n^0 \supset K_{n+1}$ *(where* K_n^0 *is the interior of* K_n*) and* $\bigcap_n K_n = K$, *and suppose that* $D(K) < \infty$. *Then we have*

$$\text{(6)} \qquad D(K) = \lim_n D(K_n).$$

Proof. — Let $T_n = T_{K_n}$, $T = T_K$, $\gamma_n = \gamma_{K_n}$, $\gamma = \gamma_K$. Then $\gamma \leq \gamma_{n+1} \leq \gamma_n$ and $\gamma \leq \beta = \lim_n \gamma_n$. By the continuity of paths, we have $X_\beta = \lim_n X_{\gamma_n}$. On the set $\{\gamma > 0\}$, we have $\{\gamma_n > 0\}$ for all n; hence $X_{\gamma_n} \in K_n$ and by continuity $X_\beta \in K$, so $\beta \leq \gamma$. Thus $\beta = \gamma$ and

$$\text{(7)} \qquad \lim_n X_{\gamma_n} = X_\gamma \quad \text{on} \quad \{\gamma > 0\}.$$

Next, on the set $\bigcap_n \{\gamma_n > 0\}$, we have

$$T_n \leq \gamma_n \leq \gamma_1 < \infty;$$

hence $\lim_n T_n \leq \infty$ and $\lim_n X(T_n) \in K$ by continuity. If $x \in E - K$, then $P^x\left\{\lim_n T_n > 0\right\} = 1$, hence

$$P^x\left\{\lim_n T_n = T\right\} = 1.$$

Thus we have for every $x \in E - K$:

$$\text{(8)} \qquad \bigcap_n \{\gamma_n > 0\} = \{\gamma > 0\},\ P^x - a.s.$$

For such an x and each bounded continuous f, we have by (7) and (8):

$$\text{(9)} \qquad \lim_n E^x\{\gamma_n > 0; f(X_{\gamma_n})\} = E^x\{\gamma > 0; f(X_\gamma)\}.$$

Let $L_n = L_{K_n}$, $L = L_K$ as given in (3). Then (9) means

that $L_n(x, .)$ converges vaguely to $L(x, .)$. Hence by assumption (a),

$$\lim_n \int_{K_1} \frac{L_n(x, dy)}{u(x, y)} = \int_{K_1} \frac{L(x, dy)}{u(x, y)}.$$

This is (6) by (4) and (5).

3. From now on we assume that u is symmetric:

$$u(x, y) = u(y, x)$$

for all (x, y). Assumption (a) above implies that $u > 0$. In order to avail ourselves of the classical theory of energy, we must assume that u is lower semi-continuous. This is assured if we strengthen (a) as follows: $u(x, y)^{-1}$ is finite continuous in (x, y).

For a compact K let M_K denote the class of signed measures with support in K, and M_K^0 the subclass of probability measures in M_K. We use the notation $U\nu$ for the function

$$U\nu(x) = \int_E u(x, y)\nu\,(dy).$$

For ν_1 and ν_2 in M_K the « mutual energy » is defined by

$$(\nu_1, \nu_2) = \int_E \int_E \nu_1\,(dx) u(x, y) \nu_2\,(dy) = \int_E (U\mu_1)\,d\mu_2$$
$$= \int_E (U\mu_2)\,d\mu_1,$$

provided that the double integral exists in the usual « absolute » sense. The quantity

$$\|\nu\|^2 = (\nu, \nu)$$

is the « energy » of ν.

The kernel u satisfies the « positivity principle » in case for any ν_1 and ν_2 in M_K we have $|(\nu_1, \nu_2)| \leqslant \|\nu_1\|\|\nu_2\|$. Then $\|\nu_1 - \nu_2\| \geqslant 0$, and $\|\nu_1 - \nu_2\| = 0$ implies

$$\|\nu_1\| = \|\nu_2\|.$$

The kernel satisfies the « energy principle » in case for any ν_1

and ν_2 in M_K, $\|\nu_1 - \nu_2\| = 0$ implies $\nu_1 = \nu_2$. Let

$$F_K^0 = \{\nu \in M_K^0 : \|\nu\|^2 < \infty\},$$

$$W(K) = \inf_{\nu \in F_K^0} \|\nu\|^2. \tag{10}$$

We have $W(K) > 0$ since $u > 0$. The lower semi-continuity of u implies the existence of a λ_K in M_K^0 such that

$$\|\lambda_K\|^2 = W(K). \tag{11}$$

Thus λ_K minimizes the energy among F_K^0. A classical argument then shows that for any $\nu \in F_K^0$, the equation

$$U\lambda_K = W(K) \tag{12}$$

holds ν-a.e. in the support of λ_K. If the kernel satisfies the « (first) maximum principle », then (12) holds ν-a.e. in E. In the standard language of potential theory, this means that (12) holds quasi-everywhere, or everywhere except for a set of (inner) capacity zero. We shall assume this in what follows. For an exposition of these results, see e.g. [3; Chapter II, § 1]. The minimization procedure indicated above is different from that used by Gauss and Frostman, but equivalent to it in effect.

Now let μ_K be the equilibrium measure for K established in [2], given by (4) above. Suppose that $\mu_K(E) > 0$, namely $D(K) < \infty$. Normalize μ_K to a probability measure σ_K by setting

$$\sigma_K = D(K)\mu_K.$$

It follows from the representation (2) that $U\mu_K \leqslant 1$ so that

$$U\sigma_K \leqslant D(K) \tag{13}$$

in E. Thus

$$\|\sigma_K\|^2 \leqslant D(K) \tag{14}$$

and $\sigma_K \in F_K^0$.

We shall call a compact K « smooth » in case every point of K is regular for K, namely,

$$P^x\{T_K = 0\} = 1$$

for every $x \in K$. Since each interior point of K is clearly regular for K, this is a condition for the boundary of K. For a smooth K the inequalities in (13) and (14) become equalities for $x \in K$ because the left member in (2) is then equal to one. Thus

$$U\sigma_K(x) = D(K) = \|\sigma_K\|^2, \quad x \in K. \tag{15}$$

It follows from (15) and (12) that

$$D(K) = \int (U\sigma_K)\, d\lambda_K = \int (U\lambda_K)\, d\sigma_K = W(K), \tag{16}$$

and so by (11)

$$\|\sigma_K\| = \|\lambda_K\|.$$

We have thus proved that for a smooth compact, the equilibrium measure obtained by the last exit method minimizes the energy. Now let K be an arbitrary compact and suppose the following is true. There exists a sequence of smooth compacts K_n such that $K_n^0 \supset K_{n+1}$ and $\bigcap_n K_n = K$. (Such a condition is often used in the study of the Dirichlet problem.) For each K_n, we have $D(K_n) = W(K_n)$, as just shown. It is clear from (10) that $W(K_n) \leqslant W(K)$ since F_K^0 increases with K. Hence it follows from (6) that

$$D(K) = \lim_n W(K_n) \leqslant W(K),$$

and so

$$\sigma_K = D(K)\mu_K \leqslant W(K)\mu_K.$$

Recalling that $U\mu_K \leqslant 1$ by (2), we have

$$U\sigma_K \leqslant W(K)U\mu_K \leqslant W(K)$$

and

$$\|\sigma_K\|^2 = \int (U\sigma_K)\, d\sigma_K \leqslant W(K). \tag{17}$$

This is the crucial inequality. We have by (12)

$$(\sigma_K, \lambda_K) = \int (U\lambda_K)\, d\sigma_K = W(K). \tag{18}$$

Using (17), (18) and (11), we obtain

$$\|\sigma_K - \lambda_K\|^2 = \|\sigma_K\|^2 - 2(\sigma_K, \lambda_K) + \|\lambda_K\|^2 \leqslant W - 2W + W = 0.$$

Thus $\|\sigma_K\| = \|\lambda_K\|$ if the positivity principle holds, and $\sigma_K = \lambda_K$ if the energy principle holds. This is what we set out to show.

4. I take this opportunity to make a correction in [2], when the paths may be discontinuous. We define, analogously to T_B and γ_B in (2):

$$T'_B = \inf\{t > 0 : X_{t-} \in B\},$$
$$\gamma'_B = \sup\{t > 0 : X_{t-} \in B\}.$$

Then in the general context treated in [2], when all paths are right continuous and have left limits, we have in place of (2) above, for every $x \in E$:

$$P^x\{T'_B < \infty\} = \int_{\bar{B}} u(x, y)\mu_B(dy) \tag{19}$$

where

$$\mu_B(A) = \int_A u(x, y)^{-1} L'_B(x, dy) \tag{20}$$

for any x and $A \in \mathscr{E}$; and

$$L'_B(x, A) = P^x\{\gamma'_B > 0;\ X(\gamma'_{B}-) \in A\}. \tag{21}$$

On p. 320 of [2], this was indicated with T_B and γ_B instead of T'_B and γ'_B. But it may happen that $X(\gamma_B) \in B$ while $X(\gamma_B-) \notin \bar{B}$, so that the measure $A \to P^x\{\gamma_B > 0;\ X(\gamma_B-) \in A\}$ need not have support in $\bar{B}$. This subtle error was discovered by John B. Walsh and led to the stated correction. The proof of (19) remains the same as well as the conclusions about the equilibrium measure and its potential. For a Hunt process, $T_B = T'_B$ a.s. for each $B \in \mathscr{E}$. Let us also remark that for the purposes of [2], we may assume that all paths are left continuous. Then of course T_B and T'_B, γ_B and γ'_B are identical. Every Hunt process, for instance, has such a left continuous version. Thus in particular for the M. Riesz potentials mentioned in [2] no change whatever is needed. Unfortunately, the method of the present note does not apply to that case because when the paths are not continuous the proof of the Lemma fails. Whether the conclusion of the Lemma, which is a necessary condition for $K \to \mu_K(E)$ to be a Choquet capacity, remains true seems in doubt.

BIBLIOGRAPHY

[1] M. BRELOT, Lectures on Potential Theory, Tata Institute, Bombay, 1960.
[2] Kai Lai CHUNG, Probability approach to the equilibrium problem in potential theory, *Ann. Inst. Fourier*, 23, 3 (1973), 313-322.
[3] N. S. LANDKOF, Foundations of Modern Potential Theory, Springer-Verlag, 1972 (English translation from Russian edition, 1966).

Manuscrit reçu le 6 mars 1975.

Kai Lai CHUNG,
Depart. of Mathematics
Stanford University
Stanford Calif. 94305 (U.S.A.).

Z. Wahrscheinlichkeitstheorie verw. Gebiete
35, 147 – 149 (1976)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

Downcrossings and Local Time

Kai Lai Chung and Richard Durrett *

Dept. of Mathematics, Stanford University, Stanford, California 94305, USA

Let $\{W(t): t \geqq 0\}$ be the standard Brownian motion with all paths continuous. Let $M(t) = \max_{0 \leqq s \leqq t} W(s)$ be the maximum process and $Y(t) = M(t) - W(t)$ be reflecting Brownian motion. If $d_\varepsilon(t)$ is the number of times Y crosses down from ε to 0 before time t, then it was Paul Lévy's idea that

$$P\{\lim_{\varepsilon \to 0} \varepsilon d_\varepsilon(t) = M(t) \text{ for all } t \geqq 0\} = 1. \tag{1}$$

In [3] Itô and McKean demonstrated the almost sure convergence of $\varepsilon d_\varepsilon(t)$ using martingale methods. To identify the limit they used the hard fact, due to Lévy, that

$$P\{\lim_{\varepsilon \to 0} (2\varepsilon)^{-1} \text{ measure } \{s: Y(s) < \varepsilon, s \leqq t\} = M(t) \text{ for all } t \geqq 0\} = 1 \tag{2}$$

and computed the second moment of the difference of the expressions in (1) and (2). In this paper, by examining the excursions in Brownian motion and using a new formula for the distribution of their maxima, we obtain a direct identification of the limit in (1) without using (2).

Let $T_x = \inf\{t: W(t) = x\}$, $T'_x = \inf\{t: Y(t) = x\}$. For $a > 0$ let $R_0^a = 0$, $R_1^a = T'_a + T'_0 \circ \theta_{T'_a}$, and for $n \geqq 2$ let $R_n^a = R_{n-1}^a + R_1^a \circ \theta_{R_{n-1}^a}$. Here $\{\theta_t, t \geqq 0\}$ is the usual collection of shift operators: $W(s, \theta_t \omega) = W(s+t, \omega)$ and if S is a random variable, $\theta_S = \theta_t$ on $\{S = t\}$. If S is a random variable, let $d_a(S) = \sup\{n: R_n^a \leqq S\}$. $d_a(S)$ is the number of downcrossings of $(0, a)$ by Y before time S.

Scaling shows that $d_{\varepsilon/m}(T_a)$ and $d_\varepsilon(T_{ma})$ have the same distribution. Using the strong Markov property $d_\varepsilon(T_{ma})$ is the sum of m independent random variables with the same distribution as $d_\varepsilon(T_a)$ so from the strong law of large numbers $\frac{\varepsilon}{m} d_{\varepsilon/m}(T_a)$ converges in probability to $E(\varepsilon d_\varepsilon(T_a))$ as $m \to \infty$.

To compute that $E(\varepsilon d_\varepsilon(T_a)) = a$ we examine the excursions in Brownian motion: (α, β) is an excursion interval of the path $Y(\cdot, \omega)$ if $\alpha < \beta$, $Y(\alpha, \omega) = 0 = Y(\beta, \omega)$ and $Y(s, \omega) > 0$ for $\alpha < s < \beta$; $\{Y(s, \omega), \alpha \leqq s \leqq \beta\}$ is called an excursion if

* Research supported in part by NSF grant GP 41710 and NSF graduate fellowship at Stanford University

(α, β) is an excursion interval—see [1] for a more complete discussion. Observe that we can count the number of down-crossings of $(0, \varepsilon)$ by Y before T_a by counting the number of excursions in $[0, T_a]$ with maxima $\geq \varepsilon$. The advantage of this viewpoint is that the excursions in $[0, T_a]$ when scaled and suitably enumerated are independent and have the same law. To state this result precisely we need to introduce the enumeration of the excursions given in [3] on page 75. Let $Z(\omega)=\{t: Y(t,\omega)=0\}$. Since Y has continuous paths, Z is a closed subset of $[0,\infty)$. Let (γ_n, β_n) be the open interval of $[0,\infty)-Z$ containing the first number of the list $1, \frac{1}{2}, \frac{3}{2}, 2, \frac{1}{4}, \frac{3}{4}, \frac{5}{4}, \frac{7}{4}, \frac{9}{4}, \frac{11}{4}, 3, \ldots$ not included in Z or $\bigcup_{m<n} (\gamma_m, \beta_m)$. Let $e_n(t)=Y(\gamma_n+t\Delta_n)/\Delta_n^{\frac{1}{2}}$ where $\Delta_n=\beta_n-\gamma_n$. Now, if we modify j_1 of (4) on page 76 of [3] to be a function of $[(Y(s), M(s));\ s\leq\gamma_1]$, then the proof on pages 75–78 gives that $\{e_n;\ n\geq 1\}$ is independent of $\{(\gamma_n,\beta_n);\ n\geq 1\}$ and M, so if we let $N_0=0$ and for $n\geq 1$, $N_n=\inf\{k>N_{n-1};\ \beta_k<T_a\}$ and define $e'_n=e_{N_n}$, then $\{e'_n: n\geq 1\}$ are independent, and each has the same law as e_1. Further, if $\Delta'_n=\beta_{N_n}-\gamma_{N_n}$, then $\{e'_n; n\geq 1\}$ and $\{\Delta'_n; n\geq 1\}$ are independent since $\{e_n; n\geq 1\}$ and $\{\Delta_n; n\geq 1\}$ are, and N_n is determined by $\{(\gamma_n,\beta_n); n\geq 1\}$ and M.

With the preliminaries on independence established, we are ready to compute the desired expectation. If we let $M'_j=\sup_{0\leq s\leq 1} e'_j(s)$, then from [1] (4.5, p. 23) or [2] (5.1, p. 21) we have

$$F(x)=P(M'_j\leq x)=1-2\sum_{n=1}^{\infty}(4n^2x^2-1)\exp(-2n^2x^2)$$

and since e'_n and Δ'_n are independent,

$$E(\varepsilon d_\varepsilon(T_a))=\varepsilon\sum_{j=1}^{\infty}\int_0^{\infty} P(M'_j>\varepsilon u^{-\frac{1}{2}})\,P(\Delta'_j\in du).$$

Now the excursion intervals $(\gamma_{N_n}, \beta_{N_n})$ correspond to jumps of the passage time process $\{T_x: x\leq a\}$, so from the Lévy decomposition ((12), p. 27 in [3]) we know that $a(2\pi u^3)^{-\frac{1}{2}}du$ is the expected number of Δ'_n with length in $(u, u+du)$, and using Fubini's theorem converts the above formula to

$$E(\varepsilon d_\varepsilon(T_a))=2a\varepsilon\int_0^{\infty}\sum_{n=1}^{\infty}\left(\frac{4n^2\varepsilon^2}{u}-1\right) exp\left(-\frac{2n^2\varepsilon^2}{u}\right)(2\pi u^3)^{-\frac{1}{2}}du.$$

Computing the above integral requires some care, because a haphazard integration term by term gives the absurdity $E(\varepsilon d_\varepsilon(T_a))=0$. However, if we integrate only on $[0, K\varepsilon^2]$, then for $n\geq K^{\frac{1}{2}}/2$ the summand in the integral is nonnegative on $[0, K\varepsilon^2]$ so we can invoke monotone convergence. To integrate the n-th term of the sum let $\alpha=4n^2\varepsilon^2$, change variables $x=\alpha/2u$ and integrate the second integral of the result by parts to get

$$\begin{aligned}2a\varepsilon\int_0^{K\varepsilon^2}\left(\frac{\alpha}{u}-1\right)e^{-\alpha/2u}(2\pi u^3)^{-\frac{1}{2}}du &= 2\alpha\varepsilon(\pi\alpha)^{-\frac{1}{2}}\int_{\alpha/2K\varepsilon^2}^{\infty}(2x^{\frac{1}{2}}-x^{-\frac{1}{2}})e^{-x}dx\\ &= a(8/\pi K)^{\frac{1}{2}}e^{-2n^2/K}.\end{aligned}$$

The remaining term

$$\int_{K\varepsilon^2}^{\infty}[1-F(\varepsilon n^{-\frac{1}{2}})](2\pi u^3)^{-\frac{1}{2}}du\leq(2/\pi K)^{\frac{1}{2}}$$

so

$$E(\varepsilon d_\varepsilon(T_a)) = \lim_{K\to\infty} a(2/\pi K)^{\frac{1}{2}} \left[1 + 2\sum_{n=1}^{\infty} e^{-2n^2/K}\right].$$

Recognizing the term in brackets as Jacobi's theta function evaluated at $2/\pi K$ and using the identity $\theta(t) = t^{-\frac{1}{2}}\theta(t^{-1})$ we get

$$E(\varepsilon d_\varepsilon(T_a)) = \lim_{K\to\infty} a\,\theta\left(\frac{\pi K}{2}\right) = a.$$

References

1. Chung, K. L.: Excursions in Brownian motion. [To appear in Arkiv för Math.]
2. Durrett, R., Iglehart, D. L.: Functionals of Brownian meander and Brownian excursion. [To appear in Annals of Probability]
3. Ito, K., McKean, H. P., Jr.: Diffusion processes and their sample paths. 2nd. ed. Berlin-Heidelberg-New York: Springer

Received October 10, 1975; in revised form November 10, 1975

Ark. Mat. 14, no. 2 (1976), 155–177.

Excursions in Brownian motion

Kai Lai Chung*

Dedicated to the Memory of Paul Lévy

1. Introduction

Paul Lévy initiated his profound study of Brownian motion on the line in his article [10] of 1939 and expounded it in one chapter of his book [11]. The article contained a wealth of ideas that inspired a generation of research. A pivot in his approach is the time set when the Brownian path takes the value zero. His idea was to use this set to partition the time axis, so as to resolve the behavior of the path into two parts: the location of the zeros, and the motion in a zero-free interval. This idea is a natural extension of the consideration of successive entrances into a fixed state in a discrete time recurrent Markov chain. But since the zeros of a Brownian path do not form a well-ordered set in the natural order of the line, the execution of the intuitive ideas is not easy. Indeed, Lévy had recourse to another time set, that when the path is surpassing its previous maximum, which he found to be of the same stochastic structure as the zero-set. He based his analysis on the new set, which also led him to the discovery of local time. Despite this brilliant detour, it turned out that a direct attack on the zeros brought quick success, as shown in Theorem 1 below. Moreover, once the crucial calculations have been made, the rest of the denouement follows the pattern of last-exit phenomenon now familiar in Markov processes. The analogy may be pushed further by treating "zero" as a unique boundary point. There is much to be gained from the analogy even from the analytic point of view. For many explicit expressions reveal themselves to be the results of juxtapositions and cancellations of basic probabilistic quantities, and their combinations and transformations are facilitated by the probabilistic insight. This is the gist of the contents of § 2, which may be regarded as a re-stumping of Lévy's old ground with a new guide.

* Research supported in part by NSF grant 41 710.

The Brownian motion restricted to the maximal zero-free interval containing the time t is called the "excursion process straddling t", and the portion of it up to t is called the "meandering process ending at t". The latter term is borrowed from D. L. Iglehart. The basic structure of meandering is given in § 3, that of excursion in § 4. It is a key fact that conditionally on the duration of each process, its law no longer depends on t, namely on its location along the time axis. By way of treating a particularly interesting functional, we derive the distribution of the maximum of each process. It is noteworthy that the result for excursion can be obtained from the corresponding distribution for the unrestricted Brownian motion by two successive partial differentiations, followed each time by setting the respective variable equal to zero. This procedure seems to mimic the action of tying down the path to vanish at both ends of the excursion. The resulting distribution may well be new and presents some analytic interest. This is pursued further in § 5. It should be clear that the method of deriving these distributions is applicable to other functionals. We content ourselves with studying occupation times in § 6, and end by extending an assertion by Lévy concerning the second moment of the occupation time of a neighborhood of zero during an excursion. This apparently fills a gap in the literature. Making use of the results in this paper it is now possible to carry further Lévy's approach to local time problems. A note on this will appear elsewhere in joint work with R. T. Durrett. [4b]. In the final § 7 we show how to obtain the clues to Brownian excursions by obvious analogy with the boundary theory for Markov chains.

Some of the results in this paper were announced in [4].

2. Basic calculations

Let $B=\{B(t),\ t\geqq 0\}$ be the standard Brownian motion with all paths continuous. For each $t>0$, we define

$$\gamma(t) = \sup\{s|s \leqq t;\ B(s) = 0\};$$

$$\beta(t) = \inf\{s|s \geqq t;\ B(s) = 0\}.$$

Then $\gamma(t)$ is the last zero of B before t, and $\beta(t)$ is the first zero of B after t. Since $B(\cdot)$ is continuous, and $P\{B(t)=0\}=0$, we have almost surely

$$\gamma(t) < t < \beta(t).$$

This is true for each t, hence also, e.g., for all rational t simultaneously. The stochastic interval $(\gamma(t), \beta(t))$ is called the *interval of excursion straddling t*. Clearly B keeps the same sign in each such interval; let $|B|=Y$, which is known as the *reflecting Brownian motion.*

The first entrance time into the singleton $\{x\}$ will be denoted by T_x,

$$T_x = \inf\{s|s > 0;\ B(s) = x\}.$$

Recall from classical theory the following formulas obtained by the reflection principle. For $t>0$ and arbitrary x,

$$P^0\{T_x \in dt\} = \frac{|x|}{\sqrt{2\pi t^3}} e^{-x^2/2t}\, dt; \tag{2.1}$$

for $t>0$, $x>0$, $y>0$,

$$P^x\{B(t)\in dy;\ T_0 > t\} = \frac{1}{\sqrt{2\pi t}}\{e^{-(x-y)^2/2t} - e^{-(x+y)^2/2t}\}\, dy. \tag{2.2}$$

Here P^x denotes the probability associated with the Brownian motion starting from x, and P^0 will be usually written as P. The differential notation above as an abbreviation for the corresponding integrated formula will be used throughout the paper.

It follows that if $0<s<t$, $x>0$, $y>0$, then we have from the meaning of $\gamma(t)$ that

$$\begin{aligned} &P\{\gamma(t) \leqq s;\ Y(s)\in dx;\ Y(t)\in dy\} = \\ &= P\{Y(s)\in dx;\ Y(u) \neq 0 \quad \text{for all} \quad u\in[s, t];\ Y(t)\in dy\} = \\ &= P\{Y(s)\in dx\} P^x\{B(t-s)\in dy;\ T_0 > t-s\} = \\ &= \sqrt{\frac{2}{\pi s}}\, e^{-x^2/2s}\, dx\, \frac{1}{\sqrt{2\pi(t-s)}}\{e^{-(x-y)^2/2(t-s)} - e^{-(x+y)^2/2(t-s)}\}\, dy. \end{aligned} \tag{2.3}$$

The first major step is to integrate out dx in the above. Straightforward calculation yields

$$\int_0^\infty \exp\left\{-\frac{x^2}{2s} - \frac{(x\pm y)^2}{2(t-s)}\right\} dt = \int_{\mp w}^\infty e^{-z^2/2}\, dz\, e^{-y^2/2t} \sqrt{\frac{s(t-s)}{t}}$$

where $w = y\sqrt{\dfrac{s}{t(t-s)}}$. Using this in (2.3) we obtain

$$P\{\gamma(t) \leqq s;\ Y(t)\in dy\} = \frac{2dy}{\pi\sqrt{t}} e^{-y^2/2t} \int_0^w e^{-z^2/2}\, dz. \tag{2.4}$$

A simple differentiation with respect to s gives the key formula below.

Theorem 1. *We have for* $0<s<t$, $y>0$,

$$P\{\gamma(t)\in ds;\ Y(t)\in dy\} = \frac{y}{\pi\sqrt{s(t-s)^3}} e^{-y^2/2(t-s)}\, ds\, dy. \tag{2.5}$$

It is trivial to integrate out dy in the above; we obtain

$$(2.6) \qquad P\{\gamma(t)\in ds\} = \frac{ds}{\pi\sqrt{s(t-s)}};$$

$$P\{\gamma(t) \leqq s\} = \frac{2}{\pi}\arcsin\sqrt{s/t}.$$

Now Theorem 1 can be cast in the following conditional form.

Corollary.

$$(2.7) \qquad P\{Y(t)\in dy\,|\,\gamma(t) = s\} = \frac{y}{t-s}e^{-y^2/2(t-s)}\,dy.$$

The last result corresponds to Lévy's Theorem 42.5 in [11] which played an essential role in his treatment. He stated it in an apparently more general form and proved it in an entirely different way. His method is to consider another process Y_1 defined by

$$Y_1(t) = \max_{0\leqq s\leqq t} B(s) - B(t),$$

and shown to be equivalent in law to Y (see [6; p. 32] for a neat proof of the latter assertion). He then used the joint distribution of $\max_{0\leqq s\leqq t} B(s)$ and $B(t)$ to derive an analogue of (2.7) for Y_1. My inability to follow his arguments was the original motivation for the present investigation. The method used here is more direct and without difficulties. In the language of point processes, where the points are the zeros of the Brownian motion, the conditioning in (2.7) is that of a "horizontal window" whereas Lévy's is a "vertical window".* As a matter of fact, since $\gamma(t)$ depends on t and is not an optional random variable, formula (2.7) by itself is not as convenient as its source (2.5).

Integrating (2.5) over s from 0 to t, we obtain

$$(2.8) \qquad \sqrt{\frac{2}{\pi t}}e^{-y^2/2t} = P\{Y(t)\in dy\}/dy = \int_0^t \sqrt{\frac{2}{\pi s}}\frac{y}{\sqrt{2\pi(t-s)^3}}e^{-y^2/2(t-s)}\,ds.$$

The identity of the first and last members above may be verified analytically, but its importance is due to the probabilistic meaning. The indicated grouping of factors in the integrand is to bring out a fundamental feature of the excursion, namely the last-exit decomposition of the Y process. To explain this and to pursue a remarkable analogy with known results for Markov chains, we introduce the appropriate nota-

* Durrett was able to justify Lévy's arguments after considerable labor.

tion below:

$$p(t;\ x, y) = \frac{1}{\sqrt{2\pi t}} e^{-(x-y)^2/2t};$$

(2.9) $$q(t;\ x, y) = p(t;\ x, y) - p(t;\ x, -y);$$

$$g(t;\ 0, y) = \frac{|y|}{\sqrt{2\pi t^3}} e^{-y^2/2t}.$$

Note $g(t; 0, y) = (\partial/\partial x)p(t; x, y)|_{x=0}$ for $y > 0$. Then (2.8) becomes, after cancelling a factor 2,

(2.10) $$p(t;\ 0, y) = \int_0^t p(s;\ 0, 0)\, g(t-s;\ 0, y)\, ds.$$

This is the last-exit (from 0) decomposition of which the analogue in Markov chains is

$$p_{0j}(t) = \int_0^t p_{00}(s)\, g_{0j}(t-s)\, ds,$$

see [1; *p.* 201]. Now observe that $p(s; 0, 0) = p(s; y, y)$, so that (2.10) may be rewritten as

(2.11) $$p(t;\ 0, y) = \int_0^t g(s; 0, y)\, p(t-s;\ y, y)\, ds.$$

In view of the interpretation in (2.1), this is just the first-entrance (into y) decomposition of which the analogue in Markov chains is

$$p_{0j}(t) = \int_0^t f_{0j}(s)\, p_{jj}(t-s)\, ds,$$

see [1; p. 205]. In Markov chains the two functions f_{0j} and g_{0j} are of course in general different. Here the fact that the same function g serves in both decompositions is another manifestation of the rich symmetry inherent in the Brownian motion. It tends however to confound matters also. The analogy with Markov chains now enables us to discover and exploit latent relations obscured by explicit expressions. First we realize that $\{g(t; 0, \cdot),\ t > 0\}$ is an *entrance law* for the *taboo transition semigroup* with $\{0\}$ as the taboo set, i.e., for $0 < s < t$ and $y > 0$, we have

$$\int_0^\infty g(s;\ 0, x)\, q(t-s;\ x, y)\, dx = g(t;\ 0, y).$$

This can again be verified analytically but its probabilistic meaning is clear. Next, initiating the procedure in Markov chains [1; p. 207], we put

(2.12) $$g(t; 0) = \int_0^\infty g(t, 0, y)\, dy = \frac{1}{\sqrt{2\pi t}};$$

(2.13) $$h(t; y) = \frac{g(t;\ 0, y)}{g(t; 0)} = \frac{y}{t} e^{-y^2/2t}$$

Rewrite (2.4) and (2.6) as follows:

$$\Gamma(s, t; dy) = P\{\gamma(t) \leqq s; Y(t) \in dy\},$$
$$\Gamma(s, t) = P\{\gamma(t) \leqq s\}.$$

Then (2.5) and (2.6) become, respectively,

$$\frac{\partial}{\partial s}\Gamma(s, t; dy) = 2p(s; 0, 0)\, g(t-s; 0, y)\, dy, \tag{2.14}$$

$$\frac{\partial}{\partial s}\Gamma(s, t) = 2p(s; 0, 0)\, g(t-s; 0). \tag{2.15}$$

The last-exit formula (2.8) may now be written as

$$P\{Y(t) \in dy\}/dy = \int_0^t h(t-s; y)\, d_s\Gamma(s, t) \tag{2.16}$$

and becomes a particular case of the more general formula, valid for $0 \leqq s_1 \leqq t$:

$$P\{\gamma(t) \leqq s_1; Y(t) \in dy\}/dy = \int_0^{s_1} h(t-s; y)\, d_s\Gamma(s, t),$$

which is the full force of (2.5).

We need another quantity to be introduced in the next proposition.

Proposition 2. *For $s > 0$ and $t > 0$, we have*

$$\int_0^\infty g(s; 0, x)\, g(t; 0, x)\, dx = \frac{1}{\sqrt{8\pi(s+t)^3}}. \tag{2.17}$$

Proof. The left member is equal to

$$\int_0^\infty \frac{1}{2\pi\sqrt{(st)^3}}\, x^2 \exp\left(-\frac{(s+t)x^2}{2st}\right) dx,$$

which is easily evaluated.

The quantity analogous to that in (2.17) for Markov chains is $\sum_j g_{0j}(s) f_{j0}(t)$. It is introduced in [2; Theorem 6.4] when 0 is regarded as a boundary point, as it might well be also in the present study. Hence we shall use a similar notation below:

$$\theta(t) = \frac{1}{\sqrt{8\pi t^3}}. \tag{2.18}$$

Indeed, we could have obtained θ from analogy with some fundamental relations in boundary theory, instead of the direct calculation in Proposition 2. This approach is given in § 7 for readers who are acquainted with boundary theory for Markov chains.

To proceed, we multiply (2.5) in the form of (2.14) by the following formula, for $0 < t < u$:

$$P\{\beta(t) \in du \mid Y(t) = y\} = P^y\{T_0 \in du - t\} = g(u-t; 0, y)\, du$$

which is obvious by the Markov property of Y. The result is

$$(2.19)\quad P\{\gamma(t)\in ds;\ Y(t)\in dy;\ \beta(t)\in du\} = = 2p(s;0,0)\,g(t-s;0,y)\,g(u-t;0,y)\,ds\,dy\,du.$$

Integrating out dy in the above by using Proposition 2, we obtain the next proposition.

Proposition 3. *For $0<s<t<u$, we have*

$$(2.20)\quad P\{\gamma(t)\in ds;\ \beta(t)\in du\} = \frac{ds\,du}{2\pi\sqrt{s(u-s)^3}} = 2p(s;0,0)\,\theta(u-s)\,ds\,du.$$

Let us introduce two more random variables:

$$(2.21)\quad L^-(t) = t-\gamma(t);\quad L(t) = \beta(t)-\gamma(t).$$

We call $(\gamma(t), t)$ the *interval of meandering ending at t*. Thus $L^-(t)$ is the duration of the meandering ending at t, whereas $L(t)$ is the duration of the excursion straddling t. For later reference we record (2.6) in the form

$$(2.22)\quad P\{L^-(t)\in dr\} = \frac{dr}{\pi\sqrt{r(t-r)}},\quad 0<r<t;$$

from which we obtain

$$(2.23)\quad P\{L(t)\in dl\,|\,L^-(t)=r\} = \frac{1}{2}\sqrt{\frac{r}{l^3}}\,dl,\quad 0<r<t\wedge l.$$

It is sometimes more convenient to use the pair $(\gamma(t), L(t))$ or $(L^-(t), L(t))$ rather than $(\gamma(t), \beta(t))$ to identify the interval of excursion straddling t, as we shall soon see.

3. The meandering process and its maximum

For each $t>0$, the process Y restricted to the interval $(\gamma(t), t)$ will be called the *meandering process ending at t*. Precisely, we define Z as follows:

$$Z(u) = Y(\gamma(t)+u)\quad\text{for}\quad 0\leqq u\leqq L^-(t).$$

For each $t>0$ and $u\geqq 0$, Z is defined only on the measurable sample set $\{u\leqq L^-(t)\}$. It would be futile to define it by decree elsewhere and we desist from doing so. The joint law of $\gamma(t)$ and the Z process is given by the next theorem.

Theorem 4. *Let $m\geqq 1$, $0<u_1<\ldots<u_m<t-s<t$, and $y_1,\ldots,y_{m+1}$ be arbitrary positive numbers. We have*

$$(3.1)\quad P\{\gamma(t)\in ds;\ Z(u_1)\in dy_1;\ \ldots;\ Z(u_m)\in dy_m;\ Y(t)\in dy_{m+1}\} =$$
$$= 2p(s;0,0)\,ds\,g(u_1;0,y_1)\,dy_1\,q(u_2-u_1;y_1,y_2)\,dy_2\ldots q(u_m-u_{m-1};y_{m-1},y_m)\,dy_m\cdot$$
$$\cdot q(t-s-u_m;y_m,y_{m+1})\,dy_{m+1}.$$

Remark. When there is no u, i.e., when $m=0$, the formula above reduces to (2.5) or (2.14).

Proof. It is sufficient to indicate the proof for $m=2$. Let $\varphi_1, \ldots, \varphi_{m+1}$ be bounded continuous functions on $(0, \infty)$; and for fixed $s>0$ let

$$d_{nk} = \frac{ks}{2^n}, \quad 0 \leqq k \leqq 2^n; \qquad I_{nk} = [d_{n,k-1}, d_{nk}), \quad 1 \leqq k \leqq 2^n.$$

If $d_{nk}+u_1<t$ and $\gamma(t)\in I_{nk}$, then $\gamma(t)=\gamma(d_{nk}+u_1)$; thus $\{\gamma(t)\in I_{nk}\}=\{\gamma(d_{nk}+u_1)\in \in I_{nk};\ \beta(d_{nk}+u_1)>t\}$. Hence we have

$$(3.2) \qquad E\{\gamma(t) < s;\ \varphi_1(Z(u_1))\varphi_2(Z(u_2))\varphi_3(Y(t))\} =$$

$$= \lim_{n\to\infty} \sum_{k=1}^{2^n} E\{\gamma(d_{nk}+u_1)\in I_{nk};\ \beta(d_{nk}+u_1) > u_2;$$

$$\varphi_1(Y(d_{nk}+u_1))\varphi_2(Y(d_{nk}+u_2))\varphi_3(Y(t))\}.$$

Here and hereafter we write $E(\Lambda; \varphi)$ for the expectation of φ over the set Λ. By Theorem 1, followed by Markov property of Y applied at $d_{nk}+u_1$, the k^{th} term above is evaluated as follows:

$$(3.3) \qquad 2\int_{I_{nk}} p(r;\, 0, 0)\, dr \int_0^\infty g(d_{nk}-r+u_1;\, 0, y)\varphi_1(y)\, \cdot$$

$$\cdot E^y\{T_0 > t-d_{nk}-u_1;\ \varphi_2(Y(u_2-u_1))\varphi_3(Y(t-d_{nk}-u_1))\}\, dy$$

where E^y is the expectation induced by P^y. For $r\in[0, s)$ let us put

$$r_n = \left(\left[\frac{2^n r}{s}\right]+1\right)\frac{s}{2^n}$$

where the square bracket denotes the greatest integer function, so that $r_n=d_{nk}$ if $r\in I_{nk}$, for $1\leqq k\leqq 2^n$. Then if we sum (3.3) over k, we obtain

$$(3.4) \qquad 2\int_0^s p(r;\, 0, 0)\, dr \int_0^\infty g(r_n-r+u_1;\, 0, y_1)\varphi_1(y_1)\, dy_1 \cdot$$

$$\cdot \int_0^\infty q(u_2-u_1;\, y_1, y_2)\varphi_2(y_2)\, dy_2 \int_0^\infty q(t-r_n-u_2;\, y_2, y_3)\varphi_3(y_3)\, dy_3.$$

Now $g(u;\, 0, y)$ and $q(u;\, x, y)$ are continuous in $u>0$ for fixed y and x, and for $0<a\leqq \leqq u\leqq b<\infty$ there is the easy domination

$$g(u;\, 0, y) \leqq \frac{|y|}{\sqrt{2\pi a}}e^{-y^2/2b}; \quad q(u;\, x, y) \leqq \frac{1}{\sqrt{2\pi a}}e^{-(x-y)^2/2b}.$$

Hence when we let $n\to\infty$ in (3.4), the result is

$$2\int_0^s p(r;0,0)\,dr\int_0^\infty g(u_1,0,y_1)\varphi_1(y_1)\,dy_1\int_0^\infty q(u_2-u_1;y_1,y_2)\varphi_2(y_2)\,dy_2\cdot$$
$$\cdot\int_0^\infty q(t-r-u_2;y_2,y_3)\varphi_3(y_3)\,dy_3.$$

Differentiation with respect to s yields the integrated form of (3.1).

Switching from $\gamma(t)$ to $L^-(t)$ and using (2.22), we obtain the conditional law below.

Corollary. *For* $0<u_1<\ldots<u_m<r<t$, *we have*

$$(3.5)\qquad P\{Z(u_1)\in dy_1;\ldots;Z(u_m)\in dy_m;Z(L^-(t))\in dy_{m+1}|L^-(t)=r\}=$$
$$=\sqrt{2\pi r}\,g(u_1;0,y_1)\,dy_1\,q(u_2-u_1;y_1,y_2)\,dy_2\ldots q(r-u_m;y_m,y_{m+1})\,dy_{m+1}.$$

It is easy to extend Theorem 4 *pro forma* to a general event belonging to the Borel field generated by the meandering process. We shall use such an extension in the specific case below. For $0<\delta<r$ and $x>0$, $\xi>0$, we have

$$(3.6)\qquad P\{\max_{\gamma(t)+\delta\le u\le t}Y(u)\le\xi;\ Y(t)\in dx|L^-(t)=r\}=$$
$$=\sqrt{2\pi r}\int_0^\xi g(\delta;0,y)P^y\{T_0>r-\delta;\ \max_{0\le u\le r-\delta}Y(u)\le\xi;Y(r-\delta)\in dx\}\,dy.$$

Let us write

$$M(t)=\max_{0\le s\le t}B(s),\quad m(t)=\min_{0\le s\le t}B(s).$$

For a fixed $\xi>0$ and $0<x\le\xi$, we put also

$$(3.7)\qquad \varphi(t;y,x)\,dx=P^y\{0<m(t)\le M(t)\le\xi;B(t)\in dx\}=$$
$$=P^0\{-y<m(t)\le M(t)\le\xi-y;B(t)\in dx-y\}.$$

Observe that the factor $P^y\{\ldots\}$ in the right member of (3.6) is then just $\varphi(r-\delta;y,x)$. It is well known that for $0<y<\xi$, $0<x<\xi$, we have

$$(3.8)\qquad \varphi(t;y,x)=\sum_{n=-\infty}^{\infty}q(t;x,y+2n\xi);$$

see, e.g., [6; p. 26]. A little inspection shows that for arbitrary $t>0$ and $x>0$, we have

$$\varphi(t;0,x)=\varphi(t;\xi,x)=0$$

and that $\varphi(t;y,x)$ is periodic in y with period 2ξ. The partial derivative of φ with respect to y is given formally by

$$(3.9)\qquad \varphi_y(t;y,x)=\frac{1}{\sqrt{2\pi t}}\sum_{n=-\infty}^{\infty}\left\{\frac{x-y-2n\xi}{t}\exp\left(-\frac{(x-y-2n\xi)^2}{2t}\right)+\right.$$
$$\left.+\frac{x+y+2n\xi}{t}\exp\left(-\frac{(x+y+2n\xi)^2}{2t}\right)\right\}.$$

The series converges uniformly in the region $0<a\leqq t\leqq b<\infty$, $0\leqq y\leqq 2\xi$, $0\leqq x\leqq 2\xi$ by easy domination. It follows that φ_y is continuous in (t, y, x) in the region $0<t<\infty$, $0\leqq y\leqq 2\xi$, $0\leqq x\leqq 2\xi$ and is indeed represented by the series. In particular, we have

$$\begin{aligned}(3.10)\qquad \varphi_y(t; 0, x) &= \sqrt{\frac{2}{\pi t}}\sum_{n=-\infty}^{\infty}\frac{x+2n\xi}{t}\exp\left(-\frac{(x+2n\xi)^2}{2t}\right) = \\ &= 2\sum_{n=0}^{\infty} g(t; 0, 2n\xi+x) - 2\sum_{n=1}^{\infty} g(t; 0, 2n\xi-x).\end{aligned}$$

We are now going to evaluate the limit of the right member of (3.6) as $\delta\downarrow 0$. By partial integration, and that $\varphi(r-\delta; 0, x)=\varphi(r-\delta, \xi, x)=0$, we have

$$\begin{aligned}(3.11)\quad \int_0^\xi g(\delta; 0, y)\varphi(r-\delta; y, x)\,dy &= \frac{1}{\sqrt{2\pi\delta}}\int_0^\xi \frac{y}{\delta}e^{-y^2/2\delta}\varphi(r-\delta; y, x)\,dy = \\ &= \frac{1}{\sqrt{2\pi\delta}}\int_0^\xi e^{-y^2/2\delta}\varphi_y(r-\delta; y, x)\,dy.\end{aligned}$$

As $\delta\downarrow 0$, it is easy to see that the last-written integral converges to $(1/2)\varphi_y(r; 0, x)$ by the continuity and boundedness of φ_y mentioned earlier. On the other hand, the left member of (3.6) converges to a similar probability with the δ there erased. Putting things together, we obtain the next theorem.

Let us put

$$M^-(t) = \max_{\gamma(t)\leqq s\leqq t} Y(s); \quad M^*(t) = \max_{\gamma(t)\leqq s\leqq \beta(t)} Y(s).$$

Thus $M^-(t)$ is the maximum of the meandering process ending at t; $M^*(t)$ is the maximum of the excursion process straddling t. The latter will be treated in § 4.

Theorem 5. *For* $0<r<t$ *and* $0<x<\xi$, *we have*

$$\begin{aligned}(3.12)\qquad & P\{M^-(t)\leqq\xi;\ Y(t)\in dx\,|\,L^-(t)=r\} = \\ &= \sqrt{2\pi r}\left\{\sum_{n=0}^{\infty} g(r; 0, 2n\xi+x) - \sum_{n=1}^{\infty} g(r; 0, 2n\xi-x)\right\}dx.\end{aligned}$$

The last expression may be written as

$$\sum_{n=-\infty}^{\infty}\frac{\partial}{\partial x}\exp\left(-\frac{(2n\xi+x)^2}{2r}\right).$$

Integrating out dx as we may term by term, we obtain

(3.13)

$$\begin{aligned}P\{M^-(t)\leqq\xi\,|\,L^-(t)=r\} &= \sum_{n=-\infty}^{\infty}\left\{\exp\left(-\frac{(2n)^2\xi^2}{2r}\right)-\exp\left(-\frac{(2n+1)^2\xi^2}{2r}\right)\right\} = \\ &= 1+2\sum_{n=1}^{\infty}(-1)^n\exp\left(-\frac{n^2\xi^2}{2r}\right).\end{aligned}$$

Multiplying (3.12) by (2.22) and integrating over r, we have

$$P\{M^-(t) \leqq \xi;\, Y(t) \in dx\} =$$
$$= \int_0^t 2p(t-r;\, 0,0)\{\textstyle\sum_{n=0}^{\infty} g(r;\, 0, 2n\xi + x) - \sum_{n=1}^{\infty} g(r;\, 0, 2n\xi - x)\}\, dr.$$

But

$$\int_0^t p(t-r;\, 0,0) g(r;\, 0,y)\, dr = p(t;\, 0,y)$$

by the last-exit decomposition (2.11); hence we obtain

$$(3.14)\quad P\{M^-(t) \leqq \xi;\, Y(t) \in dx\} = 2\textstyle\sum_{n=0}^{\infty} p(t;\, 0, 2n\xi + x) - 2\sum_{n=1}^{\infty} p(t;\, 0, 2n\xi - x) =$$
$$= 2p(t;\, 0,x) - \textstyle\sum_{n=1}^{\infty} q(t;\, x, 2n\xi).$$

Integrating out dx we get after some simplification

$$(3.15)\qquad P\{M^-(t) \leqq \xi\} = 2\textstyle\sum_{n=0}^{\infty} (-1)^n P\{n\xi \leqq B(t) < (n+1)\xi\}.$$

The last formula has also been obtained by John B. Walsh by a direct application of the reflection principle.

We can also derive (3.15) from (3.13) by an interesting detour. The key formula is as follows:

$$(3.16)\qquad \int_0^t \frac{1}{\pi\sqrt{r(t-r)}} e^{-x^2/2(t-r)}\, dr = \sqrt{\frac{2}{\pi t}} \int_x^{\infty} e^{-y^2/2t}\, dy.$$

To show this in a probabilistic setting, we rewrite the left member as

$$2\int_0^t \frac{1}{\sqrt{2\pi r}} dr \int_x^{\infty} \frac{y}{\sqrt{2\pi(t-r)^3}} e^{-y^2/2(t-r)}\, dy = 2\int_0^t p(r;\, 0,0)\, dr \int_x^{\infty} g(t-r;\, 0,y)\, dy =$$
$$= 2\int_x^{\infty} dy \int_0^t p(r;\, 0,0) g(t-r;\, 0,y)\, dr;$$

and then apply (2.11) to get

$$2\int_x^{\infty} p(t;\, 0,y)\, dy$$

which is the right member of (3.16). It follows that if we multiply (3.13) by (2.22) and then integrate with respect to r form 0 to t term by term, the result is

$$1 + 4\textstyle\sum_{n=1}^{\infty} (-1)^n \int_{n\xi}^{\infty} p(t;\, 0,y)\, dy.$$

This is seen to be the same as the right member of (3.15) by partial summation.

Let us observe that if we put $x = \xi^2/2r$ in the second member of (3.13), then we obtain the following distribution function:

$$(3.17)\qquad F(x) = \textstyle\sum_{n=-\infty}^{\infty} (-1)^n e^{-n^2 x}, \quad 0 < x < \infty.$$

This is known in the theory of theta functions. In fact, using one of Jacobi's identities (see, e.g., [12; Vol. I. p, 8]), we have

$$F(x) = \prod_{n=1}^{\infty} \frac{1-e^{-nx}}{1+e^{-nx}}, \quad 0 < x < \infty.$$

In this form it becomes clear that F is increasing and $F(0+)=0$; only $F(+\infty)=1$ is obvious from (3.17).

4. The excursion process and its maximum

For each $t>0$, the process Y restricted to the interval $(\gamma(t), \beta(t))$ is the *excursion process straddling* t. Thus it is a prolongation of the meandering process ending at t and is defined by

$$Z(u) = Y(\gamma(t)+u) \quad \text{for} \quad 0 \leqq u \leqq L(t).$$

Its fundamental structure is given in the following theorem.

Theorem 6. *Let* $m \geqq 1$, $0<u_1<\ldots<u_m<l$, *and* $y_1, \ldots, y_m$ *be arbitrary positive numbers. We have for* $s+l>t$:

$$(4.1) \qquad P\{\gamma(t)\in ds;\ Z(u_1)\in dy_1;\ \ldots;\ Z(u_m)\in dy_m;\ L(t)\in dl\} =$$
$$= 2p(s;\ 0,0)\, ds\, g(u_1;\ 0, y_1)\, dy_1\, q(u_2-u_1;\ y_1, y_2)\, dy_2 \ldots q(u_m-u_{m-1};\ y_{m-1}, y_m)\, dy_m \cdot$$
$$\cdot g(l-u_m;\ 0, y_m)\, dl.$$

Remark. When $m=1$ and when we substitute the random variable $L^-(t)$ for the constant u_1, then (4.1) reduces to (2.19). Such a substitution must of course be justified.

Proof. Again we take $m=2$ and proceed as in the proof of Theorem 4. To lighten the typographical burden we shall write

$$u_1' = d_{nk}+u_1, \quad u_2' = d_{nk}+u_2.$$

Let $T_0(u)=\inf\{t>u: B(t)=0\}$. Observe that on the set $\{u_2'<\beta(t)\}$, we have $\beta(t)=u_2'+T_0(u_2')$. Now we have, in analogy with (3.2),

$$(4.2) \qquad E\{\gamma(t) < s;\ \gamma(t)+u_2 < \beta(t);\ \varphi_1(Z(u_1))\varphi_2(Z(u_2))\varphi_3(\beta(t))\} =$$
$$= \lim_{n\to\infty} \sum_{k=1}^{2^{2n}} E\{\gamma(u_1')\in I_{nk};\ \beta(u_1') > u_2';\ \varphi_1(Y(u_1'))\varphi_2(Y(u_2'))\varphi_3(u_2'+T_0(u_2'))\}.$$

The k^{th} term in the sum is evaluated by Theorem 1 applied with $t=u_1'$, followed by Markov property of Y applied successively at u_1' and u_2'. The result is

$$2\int_{I_{nk}} p(r;\ 0,0)\, dr \int_0^\infty g(u_1'-r;\ 0, y_1)\varphi_1(y_1)\, dy_1 \int_0^\infty q(u_2'-u_1';\ y_1, y_2)\varphi_2(y_2)\cdot$$
$$\cdot E^{y_2}\{u_2'+T_0 > t;\ \varphi_3(u_2'+T_0)\}\, dy_2.$$

Using (2.1) we see that

$$E^{y_2}\{u_2'+T_0>t;\ \varphi_3(u_2'+T_0)\}=\int_{(t-u_2')^+}^{\infty} g(v-u_2';\ 0, y_2)\varphi_3(u_2'+v)\,dv.$$

When $r\in I_{nk}$, $r_n=d_{nk}$, so that $u_1'=r_n+u_1$, $u_2'=r_n+u_2$. Substituting into (4.2), we see that the sum there is equal to

$$2\int_0^s p(r;\ 0,0)\,dr\int_0^\infty g(r_n-r+u_1;\ 0, y_1)\varphi_1(y_1)\,dy_1\int_0^\infty q(u_2-u_1;\ y_1, y_2)\varphi_2(y_2)\,dy_2\cdot$$
$$\cdot\int_{(t-u_2')^+}^{\infty} g(v;\ 0, y_2)\varphi_3(u_2'+v)\,dv.$$

Letting $n\to\infty$ so that $r_n\to r$, we see that the result is tentamount to

$$P\{\gamma(t)\in ds;\ \gamma(t)+u_2<\beta(t);\ Z(u_1)\in dy_1;\ Z(u_2)\in dy_2;\ \beta(t)-\gamma(t)-u_2\in dv\}=$$
$$=2p(s;\ 0,0)\,ds\,g(u_1, 0, y_1)\,dy_1\,q(u_2-u_1;\ y_1, y_2)\,dy_2\,g(v;\ 0, y_2)\,dv,$$

if $s+u_2+v>t$; $=0$ otherwise. Writing l for u_2+v we see that this is (4.1) when $m=2$.

Corollary. *We have for* $0<s<t$; $0<u_1<\ldots<u_m<l$, *and arbitrary positive* $y_1, \ldots, y_m$:

(4.3) $$P\{Z(u_1)\in dy_1;\ \ldots;\ Z(u_m)\in dy_m|\gamma(t)=s, L(t)=l\}=$$
$$=\sqrt{8\pi l^3}\,g(u_1;\ 0, y_1)\,dy_1\,q(u_2-u_1;\ y_1, y_2)\,dy_2\ldots q(u_m-u_{m-1};\ y_{m-1}, y_m)\cdot$$
$$\cdot g(l-u_m;\ 0, y_m)\,dy_m.$$

Proof. Rewrite (2.20) as

(4.4) $$P\{\gamma(t)\in ds;\ L(t)\in dl\}=\frac{ds\,dl}{2\pi\sqrt{sl^3}}\quad\text{for}\quad t-l<s<t;$$

and divide (4.1) by (4.4). The result is (4.3).

Note that the factor $\sqrt{8\pi l^3}$ in (4.3) is just $1/\theta(l)$; so that if we put $m=1$ there and integrate out dy_1, the result indeed checks with Proposition 2.

We can now use the method of finding the distribution of $M^-(t)$ in Theorem 5 to find that of $M^*(t)$, by basing it on Theorem 6 instead of Theorem 4.

Theorem 7. *For* $0<t-s<l<\infty$, *we have*

(4.5) $$P\{M^*(t)\leqq\xi|\gamma(t)=s, L(t)=l\}=1+2\sum_{n=1}^{\infty}\left(1-\frac{4n^2\xi^2}{l}\right)\exp\left(-\frac{2n^2\xi^2}{l}\right).$$

Proof. It follows from the Corollary to Theorem 6 that for sufficiently small positive δ and ε, we have

(4.6) $$P\{\max_{s+\delta\leqq u\leqq s+l-\varepsilon} Y(u)\leqq\xi|\gamma(t)=s, L(t)=l\}=$$
$$=\sqrt{8\pi l^3}\int_0^\xi g(\delta;\ 0, y)\,dy\int_{x=0}^{\xi} P^y\{T_0>l-\delta-\varepsilon;\ \max_{0\leqq u\leqq l-\delta-\varepsilon} Y(u)\leqq$$
$$\leqq\xi;\ Y(l-\delta-\varepsilon)\in dx\}g(\varepsilon;\ 0, x).$$

Using the function φ in (3.7), the second integral above may be written as

$$\int_0^\xi \varphi(l-\delta-\varepsilon;\, y, x)\, g(\varepsilon;\, 0, x)\, dx. \tag{4.7}$$

By the argument in (3.11), we see that as $\varepsilon \downarrow 0$ this integral converges to $2^{-1}\varphi_x(l-\delta;\, y, 0)$ where φ_x is the partial derivative of $\varphi(t;\, y, x)$ with respect to x. If we substitute this limit in (4.6), its right member becomes

$$\sqrt{8\pi l^3}\,\frac{1}{2}\int_0^\xi g(\delta;\, 0, y)\,\varphi_x(l-\delta;\, y, 0)\, dy.$$

The same argument shows that as $\delta \downarrow 0$, this integral converges to

$$\sqrt{8\pi l^3}\cdot\frac{1}{4}\,\varphi_{xy}(l;\, 0, 0). \tag{4.8}$$

Here $\varphi_{xy}(t;\, y, 0)$ is the partial derivative of $\varphi_x(t;\, y, 0)$ with respect to y. Its continuity and boundedness must be checked as done before for φ_y in (3.9). From (3.10) with x and y interchanged we obtain by differentiation

$$\varphi_{xy}(t;\, y, 0) = \sqrt{\frac{2}{\pi t^3}}\sum_{n=-\infty}^{\infty}\left(1-\frac{(y+2n\xi)^2}{t}\right)\exp\left(-\frac{(y+2n\xi)^2}{2t}\right).$$

Hence the expression in (4.8) is equal to the right member of (4.5) as asserted. Theorem 8 is proved.

Putting $x=2\xi^2/l$ in the right member of (4.5), we see that the function F below,

$$F(x) = 1+2\sum_{n=1}^{\infty} e^{-n^2x}(1-2n^2x), \quad 0<x<\infty, \tag{4.9}$$

is a distribution function. Observe that $F(0+)=0$ as a consequence of Theorem 7, but this cannot be deduced by putting $x=0$ in (4.9) because the series does not converge uniformly in the neighborhood of $x=0$. Direct analytical verification of the properties F is not trivial.* One method is to pass to Laplace transforms, and using Euler's partial fraction expansion of $(e^z-1)^{-1}$ to get

$$\hat{F}(\lambda) = \int_0^\infty e^{-\lambda x}\, dF(x) = \frac{4\pi^2\lambda e^{-2\pi\sqrt{\lambda}}}{(1-e^{-2\pi\sqrt{\lambda}})^2}.$$

This shows $F(0+)=\lim_{\lambda\to\infty}\lambda\hat{F}(\lambda)=0$. But to recognize the last member as the Laplace transform of a distribution function, we need the formula, not so well known but computable:

$$E(e^{-\lambda S}) = \frac{2\pi\sqrt{\lambda}\, e^{-\pi\sqrt{\lambda}}}{1-e^{-2\pi\sqrt{\lambda}}}$$

* In fact, when the distribution was first found, the only confirmation was obtained by Iglehart on the computer.

where S is the first entrance time of the 3-dimensional Bessel process into the singleton $\{\pi/\sqrt{2}\}$. Thus F is the distribution of the sum of two independent copies of S. For a deduction which apparently goes in the opposite direction, see [13].

A quicker analytical verification of the properties of F has been furnished by W. A. Veech as follows. Introduce the theta function

$$\vartheta(x) = 1+2\sum_{n=1}^{\infty} e^{-n^2\pi x}.$$

We have then

$$F(x) = \vartheta(x)+2x\vartheta'(x).$$

Using Jacobi's functional equation

$$\vartheta(x) = \frac{1}{\sqrt{x}}\vartheta\left(\frac{1}{x}\right)$$

we obtain

$$F(\pi x) = -\frac{2}{\sqrt{x^3}}\vartheta'\left(\frac{1}{x}\right) = \frac{4\pi}{\sqrt{x^3}}\sum_{n=1}^{\infty} n^2 e^{-n^2\pi/x}. \tag{4.10}$$

If we put $\pi/x=z$, we have

$$F\left(\frac{\pi^2}{z}\right) = \frac{4}{\sqrt{\pi}} z^{3/2}\sum_{n=1}^{\infty} n^2 e^{-n^2 z}$$

and it is now clear that $\lim_{x\downarrow 0} F(x)=0$. It is trivial from (4.9) that $\lim_{x\to\infty} F(x)=1$.

Next we get from (4.9) that

$$F'(x) = 2\sum_{n=1}^{\infty} e^{-n^2 x} n^2(2n^2 x-3).$$

Hence $F'(x)>0$ for $x>3/2$. On the other hand, we get from (4.10) that

$$\pi F'(\pi x) = \frac{4\pi}{\sqrt{x^7}}\sum_{n=1}^{\infty} e^{-n^2\pi/x} n^2\left(n^2\pi-\frac{3x}{2}\right).$$

Hence $F'(x)>0$ for $0<x<2\pi^2/3$. Since $3/2<2\pi^2/3$, we have shown that $F'(x)>0$ for all $x>0$.

Iglehart [7] had the idea of studying a "scaled meandering" by the methods of weak convergence, and this was extended to a "scaled excursion" by W. D. Kaigh [9]. In the terminology used here, "scaled" means "of duration equal to one (unit)". Kaigh's result corresponding to Theorem 7 reads as follows (private communication). Let $\{X_n, n\geq 1\}$ be independent, identically distributed random variables such that X_1 takes the values $+1$ and -1 with probability $1/2$ each; and let $S_n=\sum_{k=1}^{n} X_k$; $T=\min\{n\geq 1 \mid S_n=0\}$. Then

$$\lim_{n\to\infty} P\{\max_{1\leq k\leq 2n} |S_k|/\sqrt{2n} \leq x \mid T=2n\} = F(2x^2)$$

where F is given in (4.9). A result corresponding to (3.13) is obtained when the condition "$T=2n$" above is replaced by "$T>2n$". These results can then be shown to hold for more general X_n's by an invariance principle due to Iglehart, and yield schemes which converge weakly to the scaled meandering and excursion processes. Further investigation of the various relations will appear in a paper by Durrett and Iglehart [5].

5. A curious connection

There is a way of reaching Theorem 7 via Theorem 5 which involves some interesting calculations.

Proposition 8. *For* $0<x<\xi$ *and* $u>0$ *we have*

$$(5.1) \qquad P^x\{T_0 < T_\xi;\ T_0 \in du\} = -\sum_{n=-\infty}^{\infty} p_x(u;\ 0, 2n\xi+x)\,du$$

where $p_x(u, x, y) = (\partial/\partial x)p(u;\ x, y)$.

Proof. The probability in the left member of (5.1) is expressible as

$$\lim_{\varepsilon\downarrow 0}\int_{y=0}^{\xi} P^x\{T_0 > u-\varepsilon;\ \max_{0\leq s\leq u-\varepsilon} Y(s) \leq \xi;\ Y(u-\varepsilon)\in dy\}P^y\{T_0\in d\varepsilon\} =$$

$$= \lim_{\varepsilon\downarrow 0}\int_0^{\xi} \varphi(u-\varepsilon;\ x, y)g(\varepsilon;\ 0, y)\,dy$$

in the notation of (3.7). The last limit was evaluated under (4.7) with $u=l-\delta$ and x and y interchanged as $2^{-1}\varphi_y(u;\ x, 0)$, which is seen to equal the right member of (5.1).

The Laplace transform $E^x\{T_0<T_\xi;\ e^{-\lambda T_0}\}$ is known (see. e.g., [8; p. 29]) and (5.1) may be obtained by inverting it. But the argument above, part of the proof of Theorem 7, is more in the spirit of this paper.

Now we multiply (3.12) with (5.1), and observe that the right member of (3.12) is just

$$-\sqrt{2\pi r}\sum_{n=-\infty}^{\infty} p_x(r;\ 0, 2n\xi+x).$$

We obtain thus

$$(5.2) \qquad \int_{x=0}^{\xi} P\{M^-(t) \leq \xi;\ Y(t)\in dx \mid L^-(t) = r\}P^x\{T_0 < T_\xi;\ T_0\in du\} =$$

$$= \sqrt{2\pi r}\int_0^{\xi}\sum_{n=-\infty}^{\infty} p_x(r;\ 0, 2n\xi+x)\sum_{n=-\infty}^{\infty} p_x(u;\ 0, 2n\xi+x)\,dx.$$

If we put $M^+(t) = \max_{t\leq s\leq\beta(t)} Y(s)$, then it is clear that

$$P\{M^+(t) \leq \xi;\ \beta(t)\in t+du \mid Y(t) = x\} = P^x\{T_0 < T_\xi;\ T_0\in du\}.$$

Using this in (5.2) we see that its left member is just

$$P\{M^*(t) \leq \xi;\ L(t) = r+du \mid L^-(t) = r\}.$$

Dividing this by (2.23), we get

$$(5.3)\qquad P\{M^*(t) \leqq \xi \mid L^-(t) = r,\, L(t) = r+u\} =$$
$$= \sqrt{8\pi(r+u)^3} \int_0^\xi \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} p_x(r;\, 0,\, 2m\xi+x)\, p_x(u;\, 0,\, 2n\xi+x)\, dx.$$

It is remarkable that the last-written series can be evaluated (term by term). Using the fact that p satisfies the heat equation

$$\frac{1}{2}\frac{\partial}{\partial t} p(t;\, x, y) = \frac{\partial^2}{\partial x^2} p(t;\, x, y),$$

we transform the integral by partial integration as follows:

$$\sum_m \sum_n \{p(r;\, 0,\, (2m+1)\xi)\, p_x(u;\, 0,\, (2n+1)\xi) - p(r;\, 0,\, 2m\xi)\, p_x(u;\, 0,\, 2n\xi)\} -$$
$$-2 \sum_m \sum_n \frac{\partial}{\partial u} \int_0^\xi p(r;\, 0,\, 2m\xi+x)\, p(u;\, 0,\, 2n\xi+x)\, dx$$

where all the sums range over all integers. The first double sum above vanishes because $p_x(u;\, 0, y)$ is an odd function of y. The second may be evaluated as a repeated integral by putting $k=m-n$ and summing first over n. Making use of the convolution property of p as well as its being a function of $(x-y)^2$, we reduce the sum to

$$-2\frac{\partial}{\partial u}\left\{\frac{1}{2} p(r+u;\, 0, 0) + \sum_{k=1}^{\infty} p(r+u;\, 0,\, 2k\xi)\right\}.$$

Carrying out the partial differentiation and substituting in (5.3), we see that the right member of (5.3) agrees with that of (4.5) with $l=r+u$ and $s=t-r$. This was indeed the way formula (4.5) was first "computed out". It is recorded here as an item of curiosity.

6. Occupation times

As another application of Theorem 6, we can calculate easily the expected occupation time during a meandering or excursion. Let $(a, b) \subset (0, \infty)$, and define

$$S^-(t;\, (a, b)) = \int_{\gamma(t)}^{t} I_{(a,b)}(Y(u))\, du;$$
$$S(t;\, (a, b)) = \int_{\gamma(t)}^{\beta(t)} I_{(a,b)}(Y(u))\, du;$$

where $I_{(a,b)}$ is the indicator of (a, b). We begin with the observation that for $0<t$, $x>0$, $y>0$, we have

$$(6.1)\qquad \int_0^t g(s;\, 0, x)\, g(t-s;\, 0, y)\, ds = g(t;\, 0,\, x+y).$$

This follows at once from (2.1) from the meaning of g as density of first entrance time and basic properties of the Brownian motion. Now we have by (4.3)

$$(6.2)\qquad \begin{aligned} E\{S(t;\,dx)\,|\,\gamma(t)=s,\,L(t)=l\} &= \int_0^l P\{Z(u)\in dx\,|\,\gamma(t)=s,\,L(t)=l\}\,ds \\ &= \int_0^l \sqrt{8\pi l^3}\,g(u;\,0,x)\,g(l-u;\,0,x)\,du\,dx \\ &= \sqrt{8\pi l^3}\,g(l;\,0,2x)\,dx = 4xe^{-2x^2/l}\,dx. \end{aligned}$$

This result is due to Lévy (cf. his derivation on p. 221 of [11]). It constitutes the basis of his fundamental theorem on local time cited at the end of this section.

To obtain the unconditioned expected occupation time we multiply (6.2) by (2.20) with $l=u-s$, and integrate over s:

$$(6.3)\qquad E\{S(t;\,dx)\}/dx = \frac{1}{2\pi}\int_0^t \frac{ds}{\sqrt{s}}\int_{t-s}^{\infty}\frac{4x}{\sqrt{l^3}}e^{-2x^2/l}\,dl.$$

Setting $y=2x\sqrt{(t-s)/l}$ we obtain

$$\frac{1}{x\sqrt{t-s}}\int_0^{2x} e^{-y^2/2(t-s)}\,dy$$

for the second integral in (6.3), so that

$$(6.4)\qquad E\{S(t;\,dx)\}/dx = \frac{2}{\pi}\int_0^t \frac{ds}{\sqrt{s(t-s)}}\int_0^{2x} e^{-y^2/2(t-s)}\,dy.$$

This can be evaluated by (3.16). More directly, we cast it into a probabilistic form, using (2.6) and (2.7) after integrating the latter over $(y,\,\infty)$,

$$(6.5)\qquad \begin{aligned} &2\int_0^t P\{\gamma(t)\in ds\}\int_0^{2x} P\{Y(t)>y\,|\,\gamma(t)=s\}\,dy = \\ &= 2\int_0^{2x} P\{Y(t)>y\}\,dy = 2E\{Y(t)\wedge 2x\}. \end{aligned}$$

Next, we calculate the expected occupation time during a meandering. We have by (3.5),

$$(6.6)\qquad \begin{aligned} E\{S^-(t;\,dx)/dx\,|\,L^-(t)=r\} &= \int_0^r P\{Z(u)\in dx\,|\,L^-(t)=r\}\,du \\ &= \int_0^r \sqrt{2\pi r}\,g(u;\,0,x)\,P^x\{T_0>r-u\}\,du \\ &= \sqrt{2\pi r}\int_0^r P^0\{T_x\in du\}\,P^x\{T_{2x}>r-u)\,du \\ &= \sqrt{2\pi r}\,P^0\{T_x<r<T_{2x}\}. \end{aligned}$$

Recall the notation $M(r)=\max_{0\le s\le r}B(s)$ and the basic formula

$$(6.7)\qquad P\{T_x<r\} = P\{M(r)>x\} = \sqrt{\frac{2}{\pi r}}\int_x^{\infty} e^{-y^2/2r}\,dy.$$

Using this in the last member of (6.6), we obtain

(6.8) $$E\{S^-(t;\,dx)/dx\,|\,L^-(t)=r\}=2\int_x^{2x}e^{-y^2/2r}\,dy.$$

Hence by (2.22),

$$E\{S^-(t;\,dx)\}/dx=\int_0^t\frac{2dr}{\pi\sqrt{r(t-r)}}\int_x^{2x}e^{-y^2/2r}\,dy.$$

Comparison with (6.4) and (6.5) shows that this is equal to

(6.9) $$2E\{Y(t)\wedge 2x\}-2E\{Y(t)\wedge x\}.$$

For (6.9) it is perhaps easier to calculate the difference below:

$$\begin{aligned}E\{S(t;\,dx)\}/dx-E\{S^-(t;\,dx)\}/dx&=E\Big\{\int_t^{\beta(t)}I_{(dx)}(Y(u))\,du\Big\}\Big/dx\\&=\int_0^\infty 2p(t;\,0,y)\,dy\int_0^\infty q(u;\,y,x)\,du\\&=\int_0^\infty 2p(t;\,0,y)2(x\wedge y)\,dy\\&=2E\{Y(t)\wedge x\}.\end{aligned}$$

Subtracting this from (6.5) we get (6.9).

Moments of higher order can be calculated in the same manner. For instance, we have

$$\begin{aligned}&E\{S(t;\,(a,b))^2|\gamma(t)=s,\,L(t)=l\}=\\&=2\int_0^l du_1\int_0^{l-u_1}du_2\int_a^b dx_1\int_a^b dx_2 P\{Z(u_1)\in dx_1;\,Z(u_1+u_2)\in dx_2|\gamma(t)=s,\,L(t)=l\}=\\&=2\sqrt{8\pi l^3}\int_a^b dx_1\int_a^b dx_2\int_0^l du_1\int_0^{l-u_1}du_2\,g(u_1;\,0,x_1)\cdot\\&\qquad\cdot\int_{|x_1-x_2|}^{x_1+x_2}g(u_2;\,0,z)\,dz\,g(l-u_1-u_2;\,0,x_2)\end{aligned}$$

since

$$q(u;\,x_1,x_2)=\int_{|x_1-x_2|}^{x_1+x_2}g(u;\,0,z)\,dz.$$

By (6.1) this simplifies to

$$\begin{aligned}&2\sqrt{8\pi l^3}\int_a^b dx_1\int_a^b dx_2\int_{|x_1-x_2|}^{x_1+x_2}g(l;\,0,x_1+x_2+z)\,dz=\\&=4\int_a^b dx_1\int_a^b dx_2\int_{|x_1-x_2|}^{x_1+x_2}(x_1+x_2+z)\exp\big(-(x_1+x_2+z)^2/2l\big)\,dz.\end{aligned}$$

It is possible, but perhaps futile, to evaluate this in exact terms.

In general, we have for integer $k\geqq 2$:

(6.10) $$\begin{aligned}&E\{S(t;\,(a,b))^k|\gamma(t)=s,\,L(t)=l\}=\\&=2k!\int_a^b dx_1\dots\int_a^b dx_k\int_{|x_1-x_2|}^{x_1+x_2}dz_1\dots\int_{|x_{k-1}-x_k|}^{x_{k-1}+x_k}dz_{k-1}(x_1+z_1+\dots+z_{k-1}+x_k)\\&\qquad\exp\left\{-\frac{(x_1+z_1+\dots+z_{k-1}+x_k)^2}{2l}\right\}.\end{aligned}$$

This remains true for $k=1$ by (6.2) when there is no z in the formula. A little inspection shows that

$$\int_{|x_1-x_2|}^{x_1+x_2} dz_1 \dots \int_{|x_{k-1}-x_k|}^{x_{k-1}+x_k} dz_{k-1}(x_1+z_1+\dots+z_{k-1}+x_k) \leqq (k+1)2^{k-1}x_1 \dots x_k.$$

Using this in (6.10) we obtain the inequality below.

Theorem 9. *For any $s>0$, $l>0$, $0<a<b<\infty$ and integer $k\geqq 1$, we have*

$$E\{S(t;(a,b))^k \mid \gamma(t)=s,\ L(t)=l\} \leqq (k+1)!(b^2-a^2)^k. \tag{6.11}$$

It is remarkable that the estimate does not depend on s or l. In particular, if $a=0$, $b=\varepsilon>0$, we get

$$E\{S(t;(0,\varepsilon))^k\} \leqq (k+1)!\,\varepsilon^{2k}. \tag{6.12}$$

Consequently we have for any $\lambda<1$

$$E\left\{\exp\left[\frac{\lambda S(t;(0,\varepsilon))}{\varepsilon^2}\right]\right\} < \infty. \tag{6.13}$$

On pp. 338—9 of [10], Lévy asserted an asymptotic form of (6.12) for $k=2$ (with some constant in lieu of $3!$ in front of ε^4 there), "par raison d'homogénéité". This is not clear to me. It is true that for the unconstrained Brownian motion, starting from any $x>0$, the occupation time in $(0,\varepsilon)$ until the first entrance into zero has a finite second moment. But to transport such a result to an arbitrary excursion seems to require an additional argument. This is now supplied and generalized in Theorem 9. Lévy's estimate played an essential role of the proof of his fundamental result that

$$\lim_{\varepsilon\downarrow 0} \frac{1}{2\varepsilon} \text{measure}\,\{s \mid s \leqq t;\ B(t)\in(-\varepsilon,\varepsilon)\}$$

exists almost surely for every $t>0$, and equals the local time at zero up to time t. As far as I can ascertain, no other author has returned to his original approach (see the remarks on p. 44 of [8]). For a new derivation of a related result about "downcrossings", see [4b].

It should be possible to compute from (6.10) the exact value of, say,

$$\lim_{\varepsilon\downarrow 0} E\left\{\frac{1}{\varepsilon^k} S(t;(a,a+\varepsilon))^k \mid \gamma(t)=s,\ L(t)=l\right\},$$

and thereby to determine the limit distribution of $S(t;(a,a+\varepsilon))/\varepsilon$ as $\varepsilon\downarrow 0$. The latter exists because of obvious tightness and Carleman's condition by (6.12). What it is remains to be seen.

7. Analogy with boundary theory for Markov chains

We may follow the recipe given on pp. 153—154 of [2] or pp. 85—86 of [3] to derive the basic quantities for the excursion process. First, here is the preliminary list of correspondences:

MC	a	$p_{ij}(t)$	$f_{ij}(t)$*	$l_i(t)$	e_i^a
BM	0	$p(t;\, x, y)\, dy$	$q(t;\, x, y)\, dy$	$g(t;\, 0, x)\, dx$	$1\, dx$

Apart from notations, these are obvious except possibly the last item. Now since the Brownian motion is spatially homogeneous, the Borel–Lebesgue measure on the line is invariant with respect to its transition semigroup, i.e.,

$$\int_{-\infty}^{\infty} 1\, dx\, p(t;\, x, y)\, dy = 1\, dy.$$

Hence $1dx$ plays the role of e_i^a (see p. 68 of [3]). Next ,we compute the quantity corresponding to

$$e_j^a - \sum_i e_i^a f_{ij}(t),$$

which is

$$dy - \int_0^{\infty} dx\, q(t;\, x, y)\, dy = dy - P^y\{T_0 > t\}\, dy = P^y\{T_0 \leqq t\}\, dy.$$

According to the recipe the entrance law $\{\eta_j^a(t)\}$ with respect to the minimal semigroup $\{\Phi(t)\}$ is then obtained as follows:

$$\eta_j^a(t) = \frac{d}{dt}\{e_j^a - \sum_i e_i f_{ij}(t)\}.$$

Here the corresponding step yields

$$\frac{d}{dt} P^y\{T_0 \leqq t\}\, dy = g(t;\, 0, y)\, dy.$$

Thus g is indeed the entrance law to the excursion process. Next, the formula

$$\sigma^{aa}(t) = \sum_i \eta_i^a(t) L_i^a(\infty)$$

becomes

(7.1) $$\sigma(t) = \int_0^{\infty} g(t;\, 0, x)\, 1\, dx = \frac{1}{\sqrt{2\pi t}}$$

---

* This is the transition probability for the minimal process, not the first entrance time density.

as computed in (2.12). The fundamental integral equation

(7.2) $$1 = \int_0^t E^a(ds)\sigma^{aa}(t-s), \quad 0 < t < \infty;$$

becomes

$$1 = \int_0^t E(ds)\frac{1}{\sqrt{2\pi(t-s)}}, \quad 0 < t < \infty;$$

from which we obtain the unique solution

$$e(s)\,ds = E(ds) = \sqrt{\frac{2}{\pi s}}\,ds.$$

This is just $2p(s;0,0)ds$, the probability density at zero for $Y(s)$. From a regenerative point of view this would be the fundamental quantity. Note that (7.2) turns out to be the famous arc sin law (cf. (2.6))

$$1 = \int_0^t \frac{1}{\pi\sqrt{s(t-s)}}\,ds = \frac{2}{\pi}\arcsin 1, \quad 0 < t < \infty,$$

and that there is a kind of reciprocity here: $e(t) = 2\sigma(t)$. Next, from the recipe

$$\sigma(t) = \int_t^\infty \theta(s)\,ds$$

we get from (7.1)

$$\theta(t) = -\frac{d}{dt}\sigma(t) = \frac{1}{\sqrt{8\pi t^3}}$$

as given in (2.18). After these identifications further analogy with boundary theory may be pursued easily. See for instance [4a] which contains the generic form of (2.20) above.

References

1. Chung, K. L., *Markov Chains with Stationary Transition Probabilities, second edition*, Springer-Verlag 1967.
2. Chung, K. L., On the boundary theory for Markov chains II, *Acta Math.* **115** (1966), 111—163.
3. Chung, K. L., *Lectures on Boundary Theory for Markov Chains*, Princeton University Press 1970.
4. Chung, K. L., Maxima in Brownian excursions, *Bull. Amer. Math. Soc.* **81** (1975), 742—745.
4a. Chung, K. L., A bivariate distribution in regeneration. *J. Appl. Prob.* **12** (1975), 837—839.
4b. Chung, K. L. and Durrett, R., Downcrossings and local time. To appear in *Zeitschrift für Wahrscheinlichkeitstheorie.*
5. Durrett, R. T. and Iglehart, D. L., Functionals of Brownian meandering and Brownian excursion. *To appear.*
6. Freedman, D., *Brownian Motion and Diffusion*, Holden-Day 1971.
7. Iglehart, D. L., Functional central limit theorems for random walks conditioned to stay positive, *Ann. Probability* **2** (1974), 608—619.
8. Ito, K. and McKean, Jr. H. P., *Diffusion Processes and Their Sample Paths*, Springer-Verlag, 1965.

9. Kaigh, W. D., An invariance principle for random walk conditioned by a late return to zero. *To appear in Ann. of Probability.*
10. Lévy, Paul, Sur certains processus stochastiques homogènes, *Compositio Math.* 7 (1939), 283—339.
11. Lévy, Paul, *Processus stochastiques et mouvement brownien*, second edition, Gauthier-Villars 1965 (first ed. 1948).
12. Pólya, G. and Szegő, G., *Aufgaben der Lehrsätze aus der Analysis*, Springer-Verlag 1925.
13. Williams, D., The Ito excursion law for Brownian motion. *To appear.*

Received January 7, 1976

Added in proof

Since the manuscript was prepared more than a year ago, I have given extensive lectures on it especially in Peking and Amsterdam (summer and fall of 1975), during the course of which the following amendments were made.

(1) Mr. Berber gave a quicker proof of Theorem 1 by using the equivalent Brownian motion $tB(1/t)$, thereby reducing the consideration of $(\gamma(t), Y(t))$ to $(Y(t), \beta(t))$.

(2) Besides [5], the following two papers also treat the distributions of the maxima:

D. P. Kennedy, The distribution of the maximum Brownian excursion (to appear in *J. Appl. Probability*).

D. R. Miller, The distributions of the suprema of the Brownian paths *(to appear)*.

The distribution in (4.9) was obtained by N. H. Kuiper in "Tests concerning random points on a circle". *Indag. Math.* **22** (1960), 32—37; 38—47. There it appeared as the distribution of the maximum minus the minimum in a Brownian bridge.

(3) Here is an unexpected result. If we denote the distribution in (4.9) by F_2 and that in (3.17) by F_1, then we have $F_2 = F_1 * F_1$ where $*$ denotes convolution. This is easily verified by Laplace transform and made my reference to $E(e^{-\lambda S})$ on pp. 168—169 unnecessary. The curious coincidence is still unexplained, as well as its relation to previously known results by Ito-McKean and D. Williams, concerning the path decomposition of an excursion into two Bessel (3) motions pieced together back-to-back, see § 2.10 of [8].

(4) I have calculated the first four moments of the unknown limit distribution mentioned after (6.13), namely that of $S(t; (0, \varepsilon))/\varepsilon^2$ as $\varepsilon \downarrow 0$. They are: 2, 16/3, $17 \times 2^4/15$, $31 \times 2^8/105$. The corresponding central moments are: 0, 4/3, 32/15, $69 \times 16/105$.

Received August 5, 1976

Kai Lai Chung
Department of Mathematics
Stanford University
Stanford, California 94305
U.S.A.

The Annals of Probability
1977, Vol. 5, No. 1, 82-86
with permission from Institute of Mathematical Statistics.

THE CONDENSER PROBLEM

BY K. L. CHUNG[1] AND R. K. GETOOR[2]

Stanford University and
University of California, San Diego

The condenser theorem in classical potential theory is studied within the framework of Markov processes and probabilistic potential theory. The condenser charge is expressed in terms of successive balayages of a capacitary measure.

1. Introduction. In classical potential theory on $\mathbb{R}^d$ with $d \geqq 3$ (or, more generally, in theory of Dirichlet spaces) the "condenser theorem" states the the following (see, for example, page 380 of [5]). Let G_0 and G_1 be open sets with disjoint closures $\bar{G}_0$ and $\bar{G}_1$ and assume that $\bar{G}_1$ is compact. Then there exists a potential p of a signed measure ν such that:

(i) $0 \leqq p \leqq 1$ a.e. on $\mathbb{R}^d$.

(ii) $p = 0$ a.e. on G_0 and $p = 1$ a.e. on G_1.

(iii) The support of ν^+ is contained in $\bar{G}_1$ and the support of ν^- is contained in $\bar{G}_0$.

From (i) and (ii) one would guess that $p(x)$ is just the probability that a Brownian motion starting at x hits G_1 before G_0, and consequently (i) and (ii) hold everywhere rather than almost everywhere. With this motivation it is very easy to give a probabilistic proof of the condenser theorem and to study the condenser problem within the framework of Markov processes. This note is devoted to such a study. In order to keep things simple we shall consider only Hunt processes with a locally compact metrizable state space E. (The expert should have no difficulty extending our results to the "right" processes.) Our method yields some interesting by-products. For example, it turns out that ν^+ is the capacitary measure, μ, of G_1 for the process killed when it first hits G_0 and that ν^- is the balayage of $\nu^+ = \mu$ on G_0. Moreover, we obtain an explicit formula (3.2) for μ in terms of the successive balayages on G_0 and G_1 of the capacitary measure π of G_1 for the entire process.

2. Let X be a Hunt process with state space E as in [2]. We refer the reader to [2] for all unexplained notation and terminology. Let D and B be nearly Borel sets with disjoint closures. We assume that D is transient in the sense that if $L = L_D = \sup\{t\colon X_t \in D\}$, then $L < \infty$ almost surely. (By convention the supremum of the empty set is zero and the infimum of the empty set is infinity.) As

Received March 25, 1976.

[1] Supported in part by NSF Grant GP 41710.

[2] Supported in part by NSF Grant MPS 73-04961 A01.

AMS 1970 *subject classifications.* Primary 60J25, 60J45, 60J65.

Key words and phrases. Condenser potential, balayage, Hunt processes, capacitary measure, subprocesses.

usual $T_D = \inf\{t > 0 : X_t \in D\}$ denotes the hitting time of D. Let

$$\text{(2.1)} \qquad \begin{aligned} \varphi(x) &= P_D 1(x) = P^x(T_D < \infty) = P^x(L > 0)\,, \\ p(x) &= P^x(T_D < T_B)\,. \end{aligned}$$

Then φ is an excessive function, while p is excessive relative to (X, T_B). See Section III-5 of [2]. The operators P_D and P_B are the usual balayage or hitting operators. An inclusion-exclusion argument leads to the following formula

$$p = P^\cdot(T_D < T_B) = P_D 1 - P_B P_D 1 + P_D P_B P_D 1 - \cdots.$$

The next proposition makes this precise. (C. Nevison informed us that he used it in a prior discussion.)

(2.2) Proposition. *Let* $p_n = (P_D P_B)^n P_D 1 = (P_D P_B)^n \varphi$. *Then*

$$p = \textstyle\sum_{n=0}^{\infty} (p_n - P_B p_n)\,.$$

Proof. Each p_n is excessive, bounded by one, and $P_B p_n \leqq p_n$. Therefore $0 \leqq p_n - P_B p_n \leqq 1$. Let $T_0 = 0$, $T_1 = T_D$, $T_2 = T_D + T_B \circ \theta_{T_D}, \cdots, T_{2n+1} = T_{2n} + T_D \circ \theta_{T_{2n}}$, $T_{2n+2} = T_{2n+1} + T_B \circ \theta_{T_{2n+1}}$. Thus $T_1, T_2, T_3, \cdots$ are the times of the successive visits to D, then to B, then back to D, and so on. A simple induction shows that $P_{T_{2n}} = (P_D P_B)^n$ for each $n \geqq 0$. It is straightforward to check that

$$P^x\{T_{2n+1} \leqq L \leqq T_{2n+2};\, T_D < T_B\} = p_n(x) - P_B p_n(x)$$

because L must lie in one of the intervals $[T_{2n+1}, T_{2n+2}]$. Note that the quasi-left-continuity of X implies that $\lim_n T_n = \infty$. This completes the proof of (2.2).

If $\sum p_n$ converges, then (2.2) may be written in the more agreeable form

$$\text{(2.3)} \qquad p = \textstyle\sum p_n - \sum P_B p_n\,.$$

We shall give some simple conditions that guarantee the convergence of $\sum p_n$. The hypotheses on D and B in the first paragraph of this section are still in force.

(2.4) Proposition. *Suppose there exists a nearly Borel set* G *with* $D \subset G \subset B^c$ *and satisfying*:

(i) $\sup\{U(x, G) : x \in E\} = M < \infty$.

(ii) *There exist* $t_0 > 0$ *and* $\eta > 0$ *such that* $P^x(T_{G^c} \geqq t_0) \geqq \eta$ *for all* $x \in \bar{D}^f$*—the fine closure of* D.

Then $\sum p_n(x)$ *is bounded in* x.

Proof. Let (T_n) be the sequence defined in the proof of Proposition 2.2. Then

$$p_n(x) = P_{T_{2n}} \varphi(x) = P_{T_{2n}} P_D 1(x) = P^x(T_{2n+1} < \infty)$$

for each $n \geqq 0$. Since $L < \infty$ and $T_n \uparrow \infty$ it is obvious that

$$P^x(T_{2n+1} < \infty \;\text{ for all }\; n) = 0\,.$$

Thus (2.4) is a matter of strenghtening this trivial fact to

$$\sup \textstyle\sum_n P^x(T_{2n+1} < \infty) < \infty\,.$$

If $y \in \bar{D}^f$, then by (ii)

$$E^y \int_0^{T_{G^c}} 1_G(X_t)\, dt \geqq \eta t_0 \,.$$

Now using (i) we have

$$M \geqq U(x, G) \geqq \sum_{n=0}^{\infty} E^x \int_{T_{2n+1}}^{T_{2n+2}} 1_G(X_t)\, dt$$
$$= \sum_{n=0}^{\infty} E^x\{E^{X(T_{2n+1})} \int_0^{T_B} 1_G(X_t)\, dt\} \,.$$

But $T_B \geqq T_{G^c}$ and $X(T_{2n+1}) \in \bar{D}^f$ if $T_{2n+1} < \infty$. Therefore

$$M \geqq \eta t_0 \sum_{n=0}^{\infty} P^x(T_{2n+1} < \infty) \,,$$

establishing (2.4).

REMARKS. In (2.4ii) one need only assume that $g(x) = P^x(T_{G^c} \geqq t_0) \geqq \eta$ for $x \in D$ because it is immediate from (II-4.14) of [2] that g is finely continuous. If in (2.4i) one only assumes that $U(x, G)$ is finite for each x, then the proof shows that $\sum p_n(x)$ is finite for each x.

We next formulate a simple condition under which the hypotheses of (2.4) hold. The basic result that we need is a "separation" lemma that holds when the semigroup (P_t) maps $\mathbf{C}_0$ into $\mathbf{C}_0$. Here $\mathbf{C}_0$ is the space of continuous functions on E that vanish at infinity. This result is well known and may be found in [1], for example. Nevertheless we shall give the simple proof for the convenience of the reader.

(2.5) LEMMA. *Let (P_t) map $\mathbf{C}_0$ into $\mathbf{C}_0$. Let K be compact and let G be an open neighborhood of K. Then for each $\delta > 0$ there exists a $t_0 > 0$ such that*

$$\inf_{x \in K} P^x(T_{G^c} \geqq t_0) \geqq 1 - \delta \,. \tag{2.6}$$

PROOF. We may assume without loss of generality that G has compact closure. For typographical convenience let $T = T_{G^c}$ during this proof. Since (P_t) maps $\mathbf{C}_0$ into $\mathbf{C}_0$ and $P_t f \to f$ pointwise as $t \to 0$ for each $f \in \mathbf{C}_0$, it follows that, in fact, $||P_t f - f|| \to 0$ as $t \to 0$ for each $f \in \mathbf{C}_0$ where $||\cdot||$ is the usual supremum norm. See, for example, II-(2.15) of [2]. Choose $f \in \mathbf{C}_0$ with $0 \leqq f \leqq 1$, $f = 1$ on K, and $f = 0$ on G^c. Given $\delta > 0$ there exists $t_0 > 0$ such that $||P_t f - f|| < \delta/2$ for all $t \leqq t_0$. Therefore

$$\sup_{t \leq t_0} \sup_{x \notin G} P_t f(x) < \delta/2 \tag{2.7}$$

$$\inf_{t \leq t_0} \inf_{x \in K} P_t f(x) > 1 - \delta/2 \,. \tag{2.8}$$

Thus if $x \in K$

$$1 - \delta/2 < E^x[f \circ X_{t_0}] \leqq P^x[T \geqq t_0] + E^x[f \circ X_{t_0}; T < t_0] \,, \tag{2.9}$$

and the strong Markov property implies

$$E^x[f \circ X_{t_0}; T < t_0] = E^x\{E^{X(T)}[f \circ X_{(t_0 - T)^+}]; T < t_0\} \,.$$

But $X(T) \in G^c$ if $T < \infty$ and so by (2.7) this last expectation does not exceed $\delta/2$. Combining this with (2.9) yields

$$1 - \delta/2 \leqq \inf_{x \in K} P^x(T \geqq t_0) + \delta/2 \,,$$

completing the proof of (2.5).

The following corollary is an immediate consequence of (2.4) and (2.5). Here, of course, B and D satisfy the conditions in the first paragraph of this section.

(2.10) COROLLARY. *Let (P_t) map $\mathbf{C}_0$ into $\mathbf{C}_0$ and assume that X is transient in the sense that $x \to u(x, K)$ is bounded for each compact K. Then if D has compact closure $\sum p_n(x)$ is bounded in x and*

$$p = \sum p_n - \sum P_B p_n .$$

3. In this section we shall assume that X satisfies the duality assumptions in Section VI-1 of [2] and the mild transience condition that there exists a sequence (h_n) of nonnegative functions with $h_n \uparrow 1$ and Uh_n finite for each n. Then for each x the potential kernel $u(x, y)$ is finite almost everywhere in y. See Section VI-1 of [2] for notation and terminology. As in the previous sections B and D are nearly Borel sets with disjoint closures with $L_D < \infty$. In addition throughout this section we shall suppose that the capacitary measure π_D of D exists; that is, π_D is the unique measure carried by $\bar{D}$ satisfying $\varphi = P_D 1 = U\pi_D$. For example, if $\bar{D}$ is compact and X satisfies conditions (VI-2.1), (VI-2.2), (VI-4.1), and (VI-4.2) of [2], then π_D exists. (See (VI-4.3) of [2].) However, much weaker conditions suffice. See [3] or [6] in this connection.

Let $v(x, y)$ be the potential kernel for (X, T_B)—the process X killed when it first hits B. Then v is positive kernel satisfying

$$\begin{aligned} u(x, y) &= v(x, y) + P_B u(x, y) \\ &= v(x, y) + u\hat{P}_B(x, y) . \end{aligned} \tag{3.1}$$

See [4], for example. As usual, write $V\mu(x) = \int v(x, y)\mu(dy)$ when μ is a positive measure. Let $\mu_n = \sum_{k \le n} (\hat{P}_D \hat{P}_B)^k \pi_D$, and

$$\mu_D = \lim_n \mu_n = \sum_{k=0}^{\infty} (\hat{P}_D \hat{P}_B)^k \pi_D . \tag{3.2}$$

Then μ_D is a positive measure carried by $\bar{D}$ since each μ_n is carried by $\bar{D}$. Of course, a priori, μ_D need not have any reasonable finiteness properties. However, V is a positive kernel and so

$$V\mu_D(x) = \lim V\mu_n(x)$$

exists. The fundamental identity for dual processes, VI-(1.16) of [2], yields

$$U\mu_n = \sum_{k=0}^{n} U(\hat{P}_D \hat{P}_B)^k \pi_D = \sum_{k=0}^{n} (P_D P_B)^k U\pi_D = \sum_{k=0}^{n} p_k . \tag{3.3}$$

Consequently $U\mu_n$ and $P_B U\mu_n$ are bounded for each n, and so using (2.2) and (3.1)

$$\begin{aligned} V\mu_D = \lim_n V\mu_n &= \lim_n (U\mu_n - P_B U\mu_n) \\ &= \lim_n \sum_{k=0}^{n} (p_k - P_B p_k) = p . \end{aligned}$$

Therefore

$$P^x(T_D < T_B) = p(x) = V\mu_D(x) ; \tag{3.4}$$

that is, μ_D as defined in (3.2) is the capacitary measure of D relative to the process (X, T_B).

Next suppose that $\sum p_k$ is bounded, or only finite, for each x. Conditions guaranteeing this are given in (2.4) and (2.10). Then from (3.3), $U\mu_D = \sum_{k\geq 0} p_k$ is finite and so (3.4) may be written

$$p = V\mu_D = U\mu_D - P_B U\mu_D = U\mu_D - U\hat{P}_B\mu_D = U(\mu_D - \hat{P}_B\mu_D)\,.$$

If we define $\nu = \mu_D - \hat{P}_B\mu_D$, then ν is a signed measure such that $U\nu(x) = p(x) = P^x(T_D < T_B)$. Therefore $U\nu = 1$ on D^r—the regular points of D—and 0 on B^r. But $\bar{D}$ and $\bar{B}$ are disjoint, and so $\nu^+ = \mu_D$ is carried by $\bar{D}$, more precisely by $D \cup {}^rD$ where ${}^rD$ is the set of coregular points of D, while $\nu^- = \hat{P}_B\mu_D$ is carried by $\bar{B}$, more precisely by $B \cup {}^rB$. In other words ν is the "condenser charge" for D and B and the formula

$$\nu = \mu_D - \hat{P}_B\mu_D \tag{3.5}$$

says that ν^+ is the capacitary measure μ_D of D relative to (X, T_B) and that ν^- is the balayage of $\nu^+ = \mu_D$ on B.

REMARKS. Of course, using the methods of Revuz [6], one can establish the existence of a measure μ_D such that $p = V\mu_D$ under duality and mild transience hypotheses. Then it is immediate that

$$U\mu_D = V\mu_D + P_B U\mu_D = p + U\hat{P}_B\mu_D\,. \tag{3.6}$$

But an additional "finiteness" argument seems to be necessary in order to conclude from (3.6) that

$$p = U\mu_D - U\hat{P}_B\mu_D = U(\mu_D - \hat{P}_B\mu_D)\,.$$

Our approach shows that whenever π_D exists, then μ_D exists and is given by (3.2).

REFERENCES

[1] BLUMENTHAL, R. M. (1957). An extended Markov property. *Trans. Amer. Math. Soc.* **85** 52–72.

[2] BLUMENTHAL, R. M. and GETOOR, R. K. (1968). *Markov Processes and Potential Theory.* Academic Press, New York.

[3] CHUNG, K. L. (1973). Probabilistic approach in potential theory to the equilibrium problem. *Ann. Inst. Fourier (Grenoble)* **23** 313–322.

[4] GETOOR, R. K. (1971). Multiplicative functionals of dual processes. *Ann. Inst. Fourier (Grenoble)* **21** 43–83.

[5] LANDKOF, N. S. (1972). *Foundations of Modern Potential Theory.* Springer-Verlag, Heidelberg.

[6] REVUZ, D. (1970). Mesures associées aux fonctionelles additives de Markov I and II. *Trans. Amer. Math. Soc.* **148** 501–531, and *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **16** 336–344.

DEPARTMENT OF MATHEMATICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA, SAN DIEGO
LA JOLLA, CALIFORNIA 92093

Z. Wahrscheinlichkeitstheorie verw. Gebiete
49, 237–248 (1979)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

Left Continuous Moderate Markov Processes

Kai Lai Chung* and Joseph Glover*

Stanford University, Department of Mathematics Stanford, CA 94305 USA

1. The purpose of this paper is to formulate and prove several basic results for a left continuous moderate Markov process, which are analogues of well-known results for right continuous strong Markov processes. It turns out that the first such result in our development is that about the limit or infimum of excessive functions. This was given by H. Cartan in his celebrated papers on Newtonian potentials, extended by Brelot to general potential theory, and proved by Doob by probabilistic methods. The left version of this result with certain ramifications is given in Theorems 1, 2 and 3 below. Several consequences are then drawn. In particular, Hunt's result about the regular points of a set, and Dellacherie's result on semipolar sets are given respectively in Theorems 4 and 5. Naturally, proofs of these results in the left setting follow certain well-trodden paths in the right setting, but several not so obvious detours are necessary in order to avoid the pitfalls. Some of these pitfalls are: there may be branch points; there is no zero-one law; excessive functions need not be right or left continuous on paths; the minimum of two excessive functions need not be excessive. We illustrate these pathologies by a trivial example at the end of the paper. Nevertheless, our results are as good as their right counterparts, which may or may not be surprising to the *conoscenti*. (No co-fine topology!)

Let us begin by giving a definition of a moderate Markov process. Let $(E, \mathbf{E})$ be a Lusin topological space together with its Borel field, and let $(P_t)_{t>0}$ be a Markovian semigroup on E. We set $P_0(x, \cdot) = \varepsilon_x(\cdot)$. Let $(X_t)_{t \geq 0}$ be a process with values in E, having left limits everywhere in $(0, \infty)$, defined on a measurable space $(\Omega, \mathfrak{F})$, and adapted to a filtration $(\mathfrak{F}_t)_{t \geq 0}$ with each $\mathfrak{F}_t \subset \mathfrak{F}$. We assume that $(P^x)_{x \in E}$ is a family of probability measures on $(\Omega, \mathfrak{F})$ which depend measurably on x and that $P^x\{X_0 = x\} = 1$ for each x in E. The process X is said to be a moderate Markov process with semigroup $(P_t)_{t \geq 0}$ if for each predictable stopping time T, for each positive measurable function f, and for each $t > 0$,

$$E^x\{f(X_{T+t}) | \mathfrak{F}_{T-}] = P_t f(X_{T-}) \quad \text{a.s. on } \{T < \infty\}.$$

* Research supported in part by NSF grant MCS77-01319

If X is left continuous a.s., then we may replace the right-hand side of the equation with $P_t f(X_T)$.

The class of left continuous moderate Markov processes is at least as extensive as Hunt processes. Indeed, every Hunt process (hence every Feller process) has a left continuous standard modification which is a moderate Markov process. Set $\tilde{X}_0 = X_0$, $\tilde{X}_t = X_{t-}$ for $t > 0$. Then $\tilde{X}_t$ is a left continuous process which is adapted to the filtration $(\mathfrak{F}_t)$ of the Hunt process, where $(\mathfrak{F}_t)$ is right continuous. Since $T+t$ is predictable, the quasi-left continuity of X implies that for each $t \geqq 0$, $X_{T+t} = \tilde{X}_{T+t}$ a.s. Thus in the statement of the strong Markov property for X,

$$E^x\{f(X_{T+t}) | \mathfrak{F}_T\} = P_t f(X_T),$$

we may replace X with $\tilde{X}$ to obtain

$$E^x\{f(\tilde{X}_{T+t}) | \mathfrak{F}_T\} = P_t f(\tilde{X}_T).$$

Since $P_t f(\tilde{X}_T)$ is $\mathfrak{F}_{T-}$ measurable, we may replace $\mathfrak{F}_T$ with $\mathfrak{F}_{T-}$ in the above to obtain

$$E^x\{f(\tilde{X}_{T+t}) | \mathfrak{F}_{T-}\} = P_t f(\tilde{X}_T) \quad \text{on } \{T < \infty\}.$$

Thus $\tilde{X}$ has the moderate Markov property.

2. Let $(X_t, t \geqq 0)$ be a moderate Markov process with Borelian (P_t) (where P_0 = the identity) as transition semigroup, and left continuous paths in $(0, \infty)$. By definition φ is *superaveraging* iff $\varphi \geqq 0$ and $\varphi \geqq P_t \varphi$ for every $t > 0$; and is *excessive* if in addition $\lim_{t \downarrow\downarrow 0} P_t \varphi = \varphi$. It follows then that for each $\alpha > 0$, there exists $g_n \in \mathscr{E}_+$ such that

$$\varphi = \lim_{n \to \infty} \uparrow U^\alpha g_n. \tag{1}$$

Lemma 1. *For $\alpha > 0$, $g \in b\mathscr{E}_+$, $t \to U^\alpha g(X_t)$ is left continuous.*

This is stated in [3]; here is the proof. Write

$$h(t) = e^{-\alpha t} U^\alpha g(X_t). \tag{2}$$

Then $\{h(t), t \geqq 0\}$ is a positive supermartingale. We have $\forall x$:

$$E^x\{h(t)\} = E^x\left\{\int_t^\infty e^{-\alpha s} g(X_s)\, ds\right\}. \tag{3}$$

By martingale theory, h restricted to Q (rationals) has right and left limits in $(0, \infty)$. Put

$$\varphi(t) = \lim_{Q \ni q \uparrow\uparrow t} h(q).$$

For $q < t$,

$$h(q) \geqq E^x\{h(t) | \mathfrak{F}_{q-}\};$$

hence

$$\varphi(t) \geqq E^x\{h(t)|\mathfrak{F}_{t-}\} = h(t)$$

since $h(t)$ is $\mathfrak{F}_{t-}$ measurable by left continuity of X. On the other hand, we have by bounded convergence and (3):

$$E^x\{\varphi(t)\} = \lim_{Q \ni q \uparrow \uparrow t} E^x\{h(q)\} = E^x\{h(t)\}.$$

Thus for each $t>0$:

$$P^x\{\varphi(t)=h(t)\}=1; \tag{4}$$

namely, $\{\varphi(t), t>0\}$ is a standard modification of $\{h(t), t>0\}$. By definition $t \to \varphi(t)$ is left continuous. Consider now

$$\Gamma = \{(t,w) | \varphi(t,w) \neq h(t,w)\}.$$

This is a predictable set since φ and X are left continuous. By [4; p. 72], if $P^x(\pi_\Omega \Gamma)>0$, where π_Ω is the projection on Ω, then there exists a predictable T such that $P^x\{T<\infty\}>0$, $[T] \subset \Gamma$ and so

$$\varphi(T) \neq h(T) \quad \text{ on } \{T<\infty\}. \tag{5}$$

Let $\{T_n\}$ announce T. We may take T_n to be Q-valued (see [3]). We have by (4), P^x-a.s. for all n:

$$\varphi(T_n)=h(T_n) \quad \text{ on } \{T_n<\infty\}. \tag{6}$$

By Theorem 1 of [3], we have P^x-a.s.

$$h(T_n) \to h(T) \quad \text{ on } \{T<\infty\}. \tag{7}$$

Since φ is left continuous, we have also

$$\varphi(T_n) \to \varphi(T) \quad \text{ on } \{T<\infty\}. \tag{8}$$

Thus $\varphi(T)=h(T)$, P^x-a.s. This contradicts (5) and so $P^x\{\pi_\Omega \Gamma\}=0$, for every x. Hence φ and h are indistinguishable and Lemma 1 is proved.

Lemma 2. *If φ is excessive, then a.s.*

$$t \to \varphi(X_t) \quad \textit{has right limits on } [0,\infty) \textit{ and left limits on } (0,\infty). \tag{9}$$

The same is true if φ is the pointwise limit of a sequence of excessive functions; or the infimum of such a sequence.

For the proof compare [7; p. 150].

Proof. Let g_n be as in (1), and h_n correspond to g_n as in (2). By Lemma 1, $\{h_n(t), t>0\}$ is a left continuous positive supermartingale. Let $M_n[a,b]$ denote the number of upcrossings by $h_n(\cdot)$ from $(-\infty, a)$ to $(b, +\infty)$. We have by [8; p.

128], $\forall x$:

$$E^x\{M_n[a,b]\} \leqq \frac{b}{b-a}, \tag{10}$$

and consequently if $L[a,b] = \varliminf_n M_n[a,b]$, then by Fatou's lemma

$$E^x\{L[a,b]\} \leqq \frac{b}{b-a}. \tag{11}$$

Let $M[a,b]$ denote the corresponding number of upcrossings by a path of the process $\{e^{-\alpha t}\varphi(X_t), t>0\}$. Since $e^{-\alpha t}\varphi(X_t) = \lim_{n\to\infty} h_n(t)$ for each t, trivial counting shows that

$$M[a,b] \leqq L[a,b]. \tag{12}$$

It follows from (11) and (12) and the completeness of (Ω, F, P) that

$$P^x\{M[a,b] = \infty\} = 0. \tag{13}$$

This being true for every $a<b$, the paths of $e^{-\alpha t}\varphi(X_t)$ have a.s. no oscillatory discontinuities and so (9) is true.

Next suppose $\varphi = \lim_n \varphi_n$ where each φ_n is excessive. Let M and M_n denote the number of upcrossings associated with $\varphi(X_t)$ and $\varphi_n(X_t)$, respectively. For each n, we have just proved that there is a random variable L_n such that

$$M_n[a,b] \leqq L_n[a,b],$$

and

$$E^x\{L_n[a,b]\} \leqq \frac{b}{b-a}.$$

As before, we have in our present connotation:

$$M[a,b] \leqq \varliminf_n M_n[a,b] \leqq \varliminf_n L_n[a,b].$$

It follows that (13) is again true and so the result (9) holds for φ. Finally, let $\varphi = \inf_n \varphi_n$ where each φ_n is excessive. Then $\varphi = \lim_{n\to\infty} \psi_n$ where $\psi_n = \inf_{1 \leqq m \leqq n} \varphi_m$. Since (9) is true when $\varphi = \varphi_n$, it is also true for ψ_n trivially, hence for φ as just proved. Lemma 2 is proved.

We did not prove nor need the measurability of $M[a,b]$. It would follow if we could prove that $\varphi(X_t)$ is a separable process. The same remark seems to apply to the argument in [7].

Definition. A Borel set A is thin iff $\forall x$:

$$P^x\{T_A = 0\} = 0, \qquad \text{where } T_A = \inf\{t>0 \mid X_t \in A\}. \tag{14}$$

A set is semipolar iff it is contained in a countable union of thin sets. A set is polar if $P^x\{T_A < \infty\} = 0$, $\forall x$.

For a left continuous process there is no $0-1$ law to assert that (14) is equivalent to $P^x\{T_A = 0\} < 1$.

Theorem 1. *Let f be superaveraging and $f^* = \lim_{t \downarrow \downarrow 0} P_t f$. Suppose $t \to f(X_t)$ has right and left limits on $[0, \infty)$. Then*

$$\{f > f^*\} \quad \textit{is semipolar.} \tag{15}$$

Furthermore, for each $\varepsilon > 0$:

$$\{f > f^* + \varepsilon\} \quad \textit{is thin.} \tag{16}$$

Proof. Since f^* is excessive, by Lemma 2 and our hypothesis, both limits below exist for all $t > 0$:

$$f^*(X_t)_- = \lim_{s \uparrow \uparrow t} f^*(X_s), \qquad f(X_t)_- = \lim_{s \uparrow \uparrow t} f(X_s). \tag{17}$$

Now we assume f bounded. Then we have by bounded convergence for $s \geqq 0$:

$$P_{s+} f = \lim_{t \downarrow \downarrow 0} P_{s+t} f = P_s(\lim_{t \downarrow \downarrow 0} P_t f) = P_s f^*. \tag{18}$$

Furthermore, for each x and $t > 0$:

$$\begin{aligned} E^x\{f(X_t)_-\} &= \lim_{s \uparrow \uparrow t} P_s f(x) = \lim_{s \uparrow \uparrow t} P_{s+} f(x) \\ &= \lim_{s \uparrow \uparrow t} P_s f^*(x) = E^x\{f^*(X_t)_-\} \end{aligned} \tag{19}$$

where bounded convergence is used in the first and last equation; the second equation is trivial and the third by (18). For a general superaveraging f, (19) implies that for each t and each positive constant m:

$$E^x\{(f \wedge m)(X_t)_-\} = E^x\{(f \wedge m)^*(X_t)_-\} \leqq E^x\{(f^* \wedge m)(X_t)_-\}. \tag{20}$$

But $f \geqq f^*$, so we must have equality above. Since m is arbitrary and both functions in (17) are left continuous in t, it follows that

$$P^x\{\forall\, t > 0 : f(X_t)_- = f^*(X_t)_-\} = 1. \tag{21}$$

Let $\varphi = f - f^*$, where we set $\infty - \infty = 0$, then $t \to \varphi(X_t)$ has right and left limits in $[0, \infty)$. For such a function, it is an elementary fact that

$$\{t \,|\, |\varphi(X_t) - \varphi(X_t)_-| > \varepsilon\}$$

is finite in each finite $(0, t_0)$. By (21), $\varphi(X_t)_- = 0$ for all t in $(0, \infty)$, P^x-a.s. Hence if we write

$$A_\varepsilon = \{x \,|\, \varphi(x) > \varepsilon\}.$$

we have

$$P^x\{T_{A_\varepsilon} = 0\} \leqq P^x\{X_t \in A_\varepsilon \text{ for infinitely many } t \in (0,1)\} = 0.$$

This proves (16).

We did not use Dellacherie's theorem on semipolar sets.

The following result is the generalization of the classical theorem due to Cartan, Brelot and Doob to the present setting. Let us remark that in the right continuous, strong Markov case, a very short proof was given by Chung in [2].

It is not known whether the method used there has a left-handed modification. Such a proof would be very interesting indeed. Smythe [9] gave a proof for the reverse of a right continuous, strong Markov process. Easy examples show that there are left continuous moderate Markov processes which are not such reverses. Here we give a proof for the general left case. The method reverts to Doob's old idea of supermartingale upcrossing (see Meyer [7]), but does not use Dellacherie's deep result on semipolar sets. The final results are somewhat more precise than a quick application of the latter would yield.

Theorem 2. *If f is the limit or infimum of a sequence of excessive functions, then* (15) *and* (16) *are true. Under the "hypothesis of absolute continuity"* (*Meyer's condition* (*L*)), *the conclusions remain true for the infimum of an arbitrary set of excessive functions.*

Proof. The f in the first sentence of the theorem is superaveraging, and (9) is true when $\varphi = f$ by Lemma 2. Hence the first assertion is a special case of Theorem 1. The second assertion is proved in the same way as in Meyer [7, p. 163], except for the following observation. For two superaveraging functions f and g, it is not necessarily true that $(f \wedge g)^* = f^* \wedge g^*$. But it is true that for any sequence of superaveraging f_n and any positive constant m, we have

$$(\inf_n (f_n \wedge m))^* = (\inf_n f_n)^* \wedge m \tag{22}$$

except on a semipolar set. To see this, observe that by the first assertion of Theorem 2, the left member of (22) is not smaller than the right member except on a semipolar set. On the other hand, it is not greater because for every $t > 0$, we have

$$P_t (\inf_n (f_n \wedge m)) \leqq P_t (\inf_n f_n) \wedge m.$$

Now an inspection of Meyer's proof loc. cit. shows that (22) is sufficient for the conclusions.

Lemma 3. *Let φ be excessive, then for any predictable S and T such that $S \leqq T$ we have $\forall x$:*

$$E^x\{\varphi(X_T); T < \infty \mid \mathfrak{F}_S\} \leqq \varphi(X_S) \quad \text{on } \{S < \infty\}. \tag{23}$$

Moreover, if ψ is also excessive, then

$$E^x\{(\varphi \wedge \psi)(X_T); T < \infty \mid \mathfrak{F}_S\} \leqq (\varphi \wedge \psi)(X_S) \quad \text{on } \{S < \infty\}. \tag{24}$$

Proof. By the moderate Markov property, we have for each x:

$$e^{-\alpha T} U^\alpha g(X_T) = E^x\left\{\int_T^\infty e^{-\alpha t} g(X_t)\, dt \mid \mathfrak{F}_T\right\}.$$

Hence

$$\begin{aligned} E^x\{e^{-\alpha T} U^\alpha g(X_T) \mid \mathfrak{F}_S\} &= E^x\left\{\int_T^\infty e^{-\alpha t} g(X_t)\, dt \mid \mathfrak{F}_S\right\} \\ &\leqq E^x\left\{\int_S^\infty e^{-\alpha t} g(X_t)\, dt \mid \mathfrak{F}_S\right\} = e^{-\alpha S} U^\alpha g(X_S), \end{aligned} \tag{25}$$

where we have used the convention that $\varphi(X_\infty)=0$ for any function φ. Using (25) for $g=g_n$ as in (1), we obtain (23) by first letting $n\to\infty$ and then $\alpha\downarrow 0$, involving monotone conergence both times. Now (24) is a trivial (but useful) consequence of (23) applied to φ and ψ separately.

Theorem 3. *Suppose that the f in Theorem 2 is such that $\{f>0\}$ is a set of potential zero. Then it is polar.*

Proof. Let $f=\lim\limits_n f_n$ where each f_n is excessive. The hypothesis amounts to $U^\alpha f=0$ for some $\alpha>0$. It follows that for each x, $P_t f(x)=0$ for (Lebesgue) a.e. t. Hence

$$f^*(x)=\lim_{t\downarrow\downarrow 0} P_t f(x)=0,$$

and it follows from (20) that

$$P^x\{f(X_t)_-=0 \quad \text{for all } t\in(0,\infty)\}=1. \tag{26}$$

Now $\{f(X_t), t>0\}$ is a predictable process because X is left continuous. If it is not evanescent under P^x, then by [4; p. 72] there exists a predictable T such that $P^x\{0<T<\infty\}>0$, and

$$f(X_T)>0 \quad \text{on } \{T<\infty\}. \tag{27}$$

Let $\{T_k\}$ announce T, where each T_k is predictable. Since each f_n is excessive, it follows from (24) with $\psi=m$, a constant, that

$$E^x\{f_n(X_T)\wedge m;\ T<\infty\}\leqq E^x\{f_n(X_{T_k})\wedge m;\ T_k<\infty\}. \tag{28}$$

Hence by bounded convergence we have

$$E^x\{f(X_T)\wedge m;\ T<\infty\}\leqq E^x\{f(X_{T_k})\wedge m;\ T_k<\infty\}. \tag{29}$$

Since $T_k\uparrow\uparrow T$, we have $f(X_{T_k})\to f(X_T)_-=0$ by (26); hence the right side of (29) converges to zero. But for large enough m the left side cannot be zero by (27). This contradiction proves that $f(X_t)$ is an evanescent process and so $\{f>0\}$ is polar. The proof for $f=\inf\limits_n f_n$ is similar by use of (24).

Remark. Unlike Theorem 1, Theorem 3 is not true for a superaveraging f satisfying the condition (9). Example: let b be a nonsticky boundary point in a diffusion on R^1 (or a Markov chain), and $f=1_{\{b\}}$. Then $P_t f=0$ for every $t>0$; and (9) holds when $\varphi=f$ because $\{t: X(t)=b\}$ is a discrete sequence. But $\{f>0\}=\{b\}$ is not polar.

Remark. Some of the results given above have versions in the general theory of stochastic processes. Let $(\Omega, \mathfrak{F}, P)$ be a probability space with a filtration $(\mathfrak{F}_t)$, and let M_t be a nonnegative predictable process with $E[M_0]<\infty$. Then M is said to be a predictable strong supermartingale if for any pair of predictable stopping times $S\leqq T$, $E[M_T|\mathfrak{F}_{S-}]\leqq M_S$ a.s. on $\{S<\infty\}$. Mertens [6] has proved versions of the following results for optional strong supermartingales, and his proofs apply to the predictable case with no change.

Theorem. *Let M_t be a predictable strong supermartingale.*

(i) *Then M has left limits on $(0, \infty)$.*

(ii) *If $\lim_{n\to\infty} M_{T_n} = M_T$ whenever (T_n) is a sequence of predictable stopping times announcing T, then M is left continuous. If M belongs to the class (D), this is equivalent to $\lim_{n\to\infty} E[M_{T_n}] = E[M_T]$.*

For example, if φ is an excessive funtion, choose (g_n) so that $U^\alpha g_n$ increases to φ. Since $e^{-\alpha t}\varphi(X_t) = \lim_{n\to\infty} e^{-\alpha t} U^\alpha g_n(X_t)$ is a predictable strong supermartingale, we may apply (i) above to conclude that $t \to \varphi(X_t)$ has left limits a.s. However, Mertens's techniques do not seem to yield that $t \to \varphi(X_t)$ has right limits a.s.

For an exposition of Mertens's result, see the forthcoming book by Dellacherie and Meyer, "Probabilités et potentiel", vol. 2 (Hermann, Paris).

3. For the right continuous strong Markov case, it is an essential fact that an excessive function composed with the process has right continuous sample paths. This is not the case in our situation, in general. We single out two classes of excessive functions where regularity does occur.

Propositions 1 and 2 below follow also from Merten's result (ii) in the remark above and related results.

Proposition 1. *Let f be an excessive function and T a predictable time with $P_T f(x) = f(x)$. Then $f(X_t)$ is left continuous on $]0, T]$ a.s. P^x*

Proof. Let $\Gamma_1 = \{(t,\omega): f(X_t)_- > f(X_t)\}$. If $P^x\{\pi_\Omega \Gamma_1\} > 0$, there is a predictable time S with $[S] \subset \Gamma_1$, $P^x\{S \leqq T\} > 0$. Let (S_n) be a sequence of predictable times announcing S. Then

$$f(x) = E^x\{f(X_T)\} \leqq E^x\{f(X_{S\wedge T})\} < \varliminf_{n\to\infty} E^x\{f(X_{S_n \wedge T})\} \leqq f(x)$$

by Lemma 3. Thus, $P^x\{S \leqq T\} = 0$. Now letting $\Gamma_2 = \{(t,\omega): f(X_t)_- < f(X_t)\}$ and choosing S and (S_n) as before, we have for a sufficiently large constant $R > 0$

$$E^x\{\lim_n f(X_{S_n}) \wedge R\} < E^x\{f(X_S) \wedge R\}.$$

By dominated convergence, the limit may be taken out of the expectation, and the second statement in Lemma 3 implies that

$$E^x\{f(X_S) \wedge R\} \leqq \lim_n E^x\{f(X_{S_n}) \wedge R\} < E^x\{f(X_S) \wedge R\}.$$

Thus, $P^x\{S \leqq T\} = 0$.

For Proposition 2 we assume that Ω is equipped with a family of shift operators $(\theta_t)_{t\geq 0}$ such that for each s, θ_s is a map of Ω into Ω satisfying $X_t \circ \theta_s = X_{t+s}$ a.s. for all t.

Proposition 2. *Let A be a continuous additive functional of X with potential $f(x) = E^x\{A_\infty\}$. If $f(x) < \infty$, then $f(X_t)$ is left continuous P^x a.s.*

Proof. Let Γ_1, S and (S_n) be as in Proposition 1. Then $E^x\{h(X_S)_-\} > E^x\{h(X_S)\}$. But

$$\begin{aligned} E^x\{h(X_S)_-\} &\leq \varliminf_{n\to\infty} E^x\{h(X_{S_n})\} = \varliminf_{n\to\infty} \mathrm{E}^x\{A_\infty - A_{S_n}\} \\ &= E^x\{A_\infty - A_S\} = E^x\{h(X_S)\}. \end{aligned}$$

Therefore, $P^x\{S<\infty\}=0$. Now, looking at the case for Γ_2, S and (S_n), we have that

$$E^x\{h(X_{S_n})\wedge R\}\geqq E^x\{h(X_S)\wedge R\} \qquad \text{for } R>0$$

by the second part of Lemma 3. Using dominated convergence we pass to the limit to get that $E^x\{h(X_S)_-\wedge R\}\geqq E^x\{h(X_S)\wedge R\}$. Letting R increase to ∞, we see that $P^x\{S<\infty\}=0$, and this completes the proof.

As remarked above, there is no useful $0-1$ law for moderate Markov processes. Thus, if we define $\Phi_A^\alpha(x)=E^x\{e^{-\alpha T_A}\}$, it is not apparent that the set $\{x: \Phi_A^\alpha(x)<1\}$ is semipolar. This follows, however, as a corollary of the next theorem.

Theorem 4. *Let A be Borel with* $\sup\{\Phi_A^1(x): x\in A\}=a<1$. *(Such a set is said to be totally thin.) Then $\{s<r: X_s\in A\}$ is finite a.s. for each $r>0$.*

Proof. We first show that $\{s\in[t,t+r]: X_s\in A\}$ is finite for $t>0, r>0$. Set $T_0=t$, and recursively define times $T_{n+1}=T_n+T_A\circ\theta_{T_n}$, $n=0,1,2,\ldots$. Set $R=\lim_{n\to\infty} T_n$. Then $\{R<\infty\}=\Omega_0\cup\Omega_1$, where

$$\Omega_0=\bigcup_n\{T_n=R<\infty\},$$
$$\Omega_1=\bigcap_n\{T_n<R<\infty\}.$$

Suppose $P^x(\Omega_1)>0$. Note that the (T_n) form a strictly increasing sequence on Ω_1. Choose q so large that $P^x(\Omega_1\cap\{R<q\})>P^x(\Omega_1)-\varepsilon=c>0$. Choose S_1 predictable,

$$[S_1]\subset(T_1,T_3]\cap\{(t,\omega): X_t(\omega)\in A\}\cap[0,q]$$

with $P^x\{S_1<\infty\}>c$. Set $D_1=S_1+T_A\circ\theta_{S_1}\leqq T_4$. Proceeding recursively, choose S_n predictable,

$$[S_n]\subset(T_{4n-3},T_{4n-1}]\cap\{(t,\omega): X_t(\omega)\in A\}\cap[0,q],$$

with $P^x\{S_n<\infty\}>c$. Set $D_n=S_n+T_A\circ\theta_{S_n}\leqq T_{4n}$. Then

$$\begin{aligned}ce^{-q}&\leqq E^x\{e^{-D_n}\}=E^x\{e^{-S_n}\Phi_A^1(X_{S_n})\}\\&\leqq aE^x\{e^{-S_n}\}\leqq aE^x\{e^{-D_{n-1}}\}\leqq\cdots\leqq a^n.\end{aligned}$$

Since $a<1$, and we may take n arbitrarily large, this is a contradiction. Hence $P^x(\Omega_1)=0$.

Suppose $P^x(\Omega_0)>0$. For each $\omega\in\Omega_0$, there is a sequence $(t_n(\omega))$ decreasing to $R(\omega)$ with $X_{t_n(\omega)}(\omega)\in A$. Set $S_1=(R+\varepsilon)\wedge q$ where q is chosen so large that $P^x\{(R+\varepsilon)\wedge q=R+\varepsilon\}>P^x(\Omega_0)-\delta=c>0$. Then there is an $\varepsilon_1<\varepsilon$ so that

$$P^x\{X_t(\omega)\in A \text{ for some } t\in((R+\varepsilon_1)\wedge q,S_1)\}>c.$$

Choose S_2 predictable,

$$[S_2]\subset(R\wedge q,(R+\varepsilon_1)\wedge q]\cap\{(t,\omega): X_t(\omega)\in A\},$$

with $P^x\{S_2<\infty\}>c$. Set $D_2=S_2+T_A\circ\theta_{S_2}<S_1$. Proceeding recursively, again, there is an $\varepsilon_n<\varepsilon_{n-1}$ so that

$$P^x\{X_t(\omega)\in A \text{ for some } t\in((R+\varepsilon_n)\wedge q, S_{n-1})\}>c.$$

Choose S_n predictable,

$$[S_n]\subset(R\wedge q,(R+\varepsilon_n)\wedge q]\cap\{(t,\omega)\colon X_t(\omega)\in A\},$$

with $P^x\{S_n<\infty\}>c$. Set $D_n=S_n+T_A\circ\theta_{S_n}<S_{n-1}$. Then

$$\begin{aligned} ce^{-q}&\leqq E^x\{e^{-D_2}\}=E^x\{e^{-S_2}\Phi_A^1(X_{S_2})\}\leqq aE^x\{e^{-S_2}\}\\ &\leqq aE^x\{e^{-D_3}\leqq\ldots\leqq a^nE^x\{e^{-D_{n+2}}\}.\end{aligned}$$

Since n is arbitrarily large, we conclude as before that $P^x(\Omega_0)=0$.

It remains to drop the hypothesis that $t>0$. This amounts to proving that $P^x\{T_A=0\}=0$, which is similar to the proof that $P^x(\Omega_0)=0$.

Corollary 1. *Let A be a Borel set with $\Phi_A^1(x)<1$ on A. Then A is semipolar. In particular, if B is an arbitrary Borel set and $B^r=\{x\colon \Phi_B^1(x)=1\}$, then $B-B^r$ is semipolar.*

Proof. Let $A_n=A\cap\left\{\Phi_A^1(x)\leqq 1-\frac{1}{n}\right\}$. Then $A=\bigcup_n A_n$ and $\Phi_{A_n}^1\leqq\Phi_A^1$. Therefore, $\sup\{\Phi_{A_n}^1(x)\colon x\in A_n\}\leqq 1-\frac{1}{n}$. Applying Theorem 4, it follows that A is semipolar.

If $x\in B-B^r$, then $\Phi_{B-B^r}^1(x)\leqq\Phi_B^1(x)<1$. Thus $B-B^r$ is semipolar. This is Hunt's theorem (see [7, p. 148]).

For the remainder of this note, we fix X a left continuous moderate Markov process satisfying Meyer's hypothesis (L). Let λ be a reference probability measure for X (see [7, p. 158]). Recall that

$$\text{if } f \text{ is excessive and } \lambda(f)=0, \text{ then } f=0. \tag{30}$$

The proof of the following result requires no change in this situation.

Proposition 3. *Let μ be a measure on E. Then μ may be decomposed as $\mu=\mu_1+\mu_2$ where μ_1 does not charge any semipolar set and μ_2 is carried by a semipolar set.*

Using Theorems 2 and 4, the proof of the following result is valid here [7, p. 180]. Recall that a set P is finely perfect if $P=\{\Phi_P^1=1\}$.

Proposition 4. *Let A be compact. Then there exists a finely perfect set $P\subset A$ such that $A-P$ is semipolar.*

We now extend Dellacherie's proof of his characterization of semipolar sets to this situation.

Theorem 5. *Let X_t be a left continuous moderate Markov process with fundamental reference probability measure λ. Let G be Borel with $P^\lambda\{X_t\in G$ at most countably often$\}=1$. Then G is semipolar.*

Proof. We sketch Dellacherie's argument here (see [5, p. 112]). Set $\Gamma=\{(t,\omega)$: $X_t(\omega)\in G\}$. By [4, VI-T33], $\Gamma\subset\bigcup_n [T_n]$ where (T_n) is a sequence of predictable stopping times. Define a measure μ on E by setting

$$\mu(H)=E^\lambda\left\{\sum_{n=1}^{\infty} 2^{-n} 1_H\circ X_{T_n}\right\}.$$

If $H\subset G$ and $\mu(H)=0$, then H is polar by (30). By Proposition 3, $\mu=\mu_1+\mu_2$, where μ_1 does not charge any semipolar set, and μ_2 is carried by a semipolar set. We show $\mu_1\equiv 0$. It suffices to show that every compact $K\subset G$ is semipolar. But K may be written as $P\cup(K-P)$ where P is a finely perfect set and $K-P$ is semipolar by Proposition 4. We prove below in Theorem 8 that $P=\emptyset$. Thus $\mu=\mu_2$, and μ_2 is therefore carried by a semipolar set $L\subset G$. But $\mu(G-L)=0$ implies $G-L$ is polar. Thus G is semipolar.

The next proof is a modification of one given in [7, p. 182].

Lemma 4. *Let P be finely perfect, and satisfy the hypothesis of Theorem 5. Then $P=\emptyset$.*

Proof. $\Gamma=\{(t,\omega)\colon X_t(\omega)\in P\}\subset\bigcup_n [T_n]$ is predictable.

$$P^\lambda\{X_{T_n}\in P;\ T_P\circ\theta_{T_n}=0\}=E^\lambda\{X_{T_n}\in P;\ P^{X_{T_n}}[T_P=0]\}=P^\lambda\{X_{T_n}\in P\}.$$

Thus $T_n(\omega)$, whenever $X_{T_n}\in P$, is a limit from the right of times $(t_n(\omega))$ with $X_{t_n(\omega)}(\omega)\in P$. Set $\Gamma(\omega)=\{t\colon X_t(\omega)\in P\}$. It follows from the preceding sentence that $\bar\Gamma(\omega)$ has no isolated point and hence is perfect and therefore uncountable. We show $\bar\Gamma(\omega)-\Gamma(\omega)$ is countable. Recall $\Gamma(\omega)=\{t\colon \Phi_P^1(X_t(\omega))=1\}$. Now $t\in\bar\Gamma(\omega)-\Gamma(\omega)$ exactly when there exist $t_k(\omega)\to t(\omega)$ such that $\Phi_P^1(X_{t_k(\omega)}(\omega))=1$ and $\Phi_P^1(X_{t(\omega)}(\omega))\neq 1$. But $e^{-t}\Phi_P^1(X_t)$ has only finitely many upcrossings over any level (a,b) by Lemma 2. If there were an uncountable number of points in $\bar\Gamma(\omega)-\Gamma(\omega)$, there would exist $0<a<b<1$ such that $e^{-t}\Phi_P^1(X_t)$ had an infinite number of upcrossings over (a,b), which is impossible. Thus $\Gamma(\omega)$ is a.s. uncountable, which contradicts $X_t\in G$ only countably often. Therefore, $P=\emptyset$.

We conclude with an example which, although trivial, illustrates much of the pathology associated with moderate Markov processes.

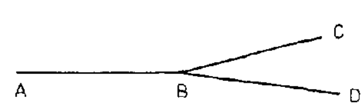


Fig. 1

Let X_t be the process uniform motion to the right on the state space given in Fig. 1. Upon reaching B, the process moves toward C with probability $\frac{1}{2}$ and moes toward D with probability $\frac{1}{2}$. Then X_t is a normal continuous process. However, the strong Markov property fails to hold at B because the $0-1$ law does not hold for the hitting time of $(B,C]$. Let $[A,B]$ (resp. $(B,C]$) denote the points between A and B,

including A and B (resp. the points between B and C, including C and excluding B). Then $P^B[T_{(B,C]}=0]=\frac{1}{2}$. Thus the $0-1$ law does not hold at B for this hitting time.

If we let

$$f(x)=\begin{cases} 1 & \text{on } [A,B) \\ \frac{1}{2} & \text{on } B \\ 1 & \text{on } (B,C] \\ 0 & \text{on } (B,D] \end{cases}$$

then $f(x)$ is excessive, but $t\to f(X_t)$ is neither right or left continuous. Moreover, $f\wedge \frac{1}{2}$ is not excessive.

Finally, we give a trivial example of a continuous moderate Markov process with state space given in Fig. 2 which is not the reverse of a strong Markov process.

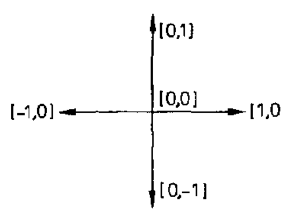


Fig. 2

On $((0,-1),(0,0))$ the process is uniform motion up; on $((-1,0),(0,0))$ the process is uniform motion to the right. At $(0,0)$, the process proceeds up toward $(0,1)$ with probability $\frac{1}{2}$ and toward $(1,0)$ with probability $\frac{1}{2}$. This process is moderate but not strong Markov, and so is the reverse by symmetry of the state space.

References

1. Chung, K.L., Doob, J.L.: Fields, optionality, and measurability. Amer. J. Math. **87**, 397–424 (1965)
2. Chung, K.L.: A simple proof of Doob's convergence theorem, Sem. de Probabilités V. Lecture Notes in Math. **191**. Berlin-Heidelberg-New York: Springer 1970
3. Chung, K.L.: On the fundamental hypotheses of Hunt processes. Symposia Mathematica, Istituto Nazionale di Alta Matematica, Vol. IX, 43–52. New York-London: Academic Press 1972
4. Dellacherie, C.: Capacités et processus stochastiques. Ergebnisse der Mathematik, Band **67**. Berlin-Heidelberg-New York: Springer 1972
5. Dellacherie, C.: Ensembles aléatoires, Sem. de Probabilités III. Lecture Notes in Math. **88**. Berlin-Heidelberg-New York: Springer 1969
6. Mertens, J.F.: Théorie des processus stochastiques généraux applications aux surmartingales. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **22**, 45–68 (1972)
7. Meyer, P.A.: Processus de Markov. Lecture Notes in Math. **26**. Berlin-Heidelberg-New York: Springer 1967
8. Meyer, P.A.: Probabilités et potentiel. Paris: Hermann 1966
9. Smythe, R.T.: Remarks on the hypotheses of duality, Sem. de Probabilités VIII. Lecture Notes in Mathematics **381**. Berlin-Heidelberg-New York: Springer 1972/1973

Received December 5, 1978

PROBABILITY
AND
MATHEMATICAL STATISTICS

Vol. 1, Fasc. 2 (1980), p. 99-108

EQUILIBRIUM AND ENERGY

BY

KAI LAI CHUNG AND MURALI RAO (STANFORD, CALIFORNIA)

Abstract. In this paper* it is shown that the equilibrium measure ν for a compact K in potential theory can be related with a unique invariant measure π for a discrete time Markov process by the formula $\pi(dy) = \varphi(y)\nu(dy)$. The chain has the transition function $L(x, A)$, where L is the last-exit kernel in [1]. For a general non-symmetric potential density u the modified energy $I(\lambda) = \iint \lambda(dx)u(x, y)\varphi(y)^{-1}\lambda(dy)$ and the Gauss quadratic $G(\lambda) = I(\lambda)-2\lambda(K)$ are introduced. Then G is minimized by π among all signed measures λ on K of finite modified energy, provided I is positive. This includes the classical symmetric case of Newtonian and M. Riesz potentials as a special case. The modification corresponds to a time change for the underlying Markov process. The positivity of I is established for a class of signed measures associated with continuous additive functionals in the sense of Revuz.

Introduction. In electrostatics, the equilibrium charge on a conductor minimizes the potential energy. Gauss showed this but assumed the existence of a minimum, which assumption became known as the Dirichlet principle. The method was extended by Frostman to M. Riesz potentials. More generally, a theory of energy has been developed for symmetric potential kernels (see, e.g., [6]). From quite another direction the existence of the equilibrium measure was established in [1] by modern methods of Markov processes, together with its probabilistic significance in terms of a last-exit distribution. The question arises whether such a measure minimizes the corresponding energy. For a symmetric kernel this was answered in the affirmative in [2]

* The research of the first-named author was supported in part by a Guggenheim Fellowship in 1976-1977, and in part by NSF Grant MCS 77-01319-A02. The research of the second-named author was supported by NSF Grant MCS 77-01319-A02.

under further superfluous conditions, but a simpler proof due to J. R. Baxter is contained in Section 3. Here we prove a general minimization result for a *modified* energy which corresponds to a time change of the process. The significance of this is not clear yet, but it contains the classical symmetric case without modification. Since energy concepts in the non-symmetric case have been little studied, we hope the results presented in the sequel may spur on further research in this direction. Let us add that although the term "equilibrium" is often used to connote a "steady state" in physics, namely a stationary or invariant distribution, the association of the electrostatic equilibrium measure with the invariant distribution of a simple Markov chain described in Section 1 is apparently new. What is its physical significance?

1. Let U be the potential kernel of a Markov process and let u be its density with respect to some reference measure m such that

$$U(x, A) = \int_A u(x, y)\, m(dy)$$

for each Borel set A. Let K be a compact set and suppose there exists a measure ν with support in K such that

$$1 = \int_K u(x, y)\, \nu(dy) \quad \text{for all } x \in K. \tag{1.1}$$

Then ν is called the *equilibrium measure* for K (it is unique under general conditions). We introduce the kernel L as

$$L(x, A) = \int_A u(x, y)\, \nu(dy), \qquad x \in K,\ A \in \mathfrak{B}(K), \tag{1.2}$$

where $\mathfrak{B}(K)$ is the Borel field of K. Then (1.1) takes the form

$$L(x, K) = 1 \quad \text{for all } x \in K. \tag{1.3}$$

Thus L is strictly Markovian and a (discrete time) Markovian chain with state space K may be constructed from L.

Under the basic assumptions of [1] and [3], the measure ν exists for each compact K provided all points of K are regular for K. In general, the constant 1 in (1.1) is to be replaced by $P_K 1$, the hitting probability of K. Furthermore, the method shows that L is the "last-exit kernel". Although we are particularly interested in the setting of [1] and [3], we shall here simply assume the validity of (1.1) without reference to the specific conditions under which it is derived. Other conditions on the function u needed for further development will be added as we proceed.

The following proposition is due to Mamoru Kanda, who improved an easier, less satisfactory condition:

Proposition 1. *If for each fixed y the function $u(\cdot, y)$ is lower semi-continuous, then the set of functions $\{u(x, \cdot), x \in K\}$ is uniformly integrable with respect to ν.*

Proof. Let $0 \leqslant f \leqslant 1$ and let f be Borelian. Then by (1.3) we have

$$\int_K u(x, y) f(y) \nu(dy) + \int_K u(x, y)[1-f(y)]\nu(dy) = 1 \quad \text{for all } x.$$

Both terms on the left-hand side are lower semi-continuous, hence both are continuous in x. Let $A_n \subset K$ and $\nu(A_n)\downarrow 0$. Then, by Dini's theorem,

$$\int_{A_n} u(x, y)\nu(dy)\downarrow 0$$

uniformly for $x \in K$. This together with (1.3) establishes the asserted uniform integrability.

Proposition 2. *Under the condition of Proposition 1 there exists a unique probability measure π on K such that $\pi = \pi L$, namely,*

$$\pi(A) = \int_K \pi(dx) L(x, A), \qquad A \in \mathfrak{B}(K). \tag{1.4}$$

π is absolutely continuous with respect to ν.

Proof. The existence follows at once from an old theorem due to Doblin since uniform integrability is much stronger than his hypothesis (D) (see [5], p. 192). Alternatively, we can apply Schauder's fixed point theorem as follows. Consider the class $M(K)$ of probability measures on K. This is convex and compact with respect to the vague topology. The kernel L in (1.2) induces a mapping $\lambda \to \lambda L$ of K into K, where

$$\lambda L(\cdot) = \int \lambda(dx) L(x, \cdot). \tag{1.5}$$

If $\lambda_n \to \lambda$ vaguely, then for each $f \in C(K)$ we have

$$\lambda_n L(f) = \int \lambda_n(dx) L(x, f) \to \int \lambda(dx) L(x, f)$$

because $x \to L(x, f)$ is continuous, as shown in the proof of Proposition 1 (even for a bounded Borelian f). Thus there exists a fixed point under the mapping which is the π in (1.4).

If we put

$$\varphi(y) = \int \pi(dx) u(x, y). \tag{1.6}$$

then

$$\pi(dy) = \varphi(y)\nu(dy). \tag{1.7}$$

Substituting back in (1.6), we obtain

$$\varphi(y) = \int \nu(dx)\varphi(x) u(x, y). \tag{1.8}$$

To prove the uniqueness, suppose that π_1 is another probability measure such that $\pi_1 = \pi_1 L$, and put $\mu = \pi - \pi_1$. Using (1.7), (1.8) and their analogues for π_1, we have

$$\varphi(y) - \varphi_1(y) = \int \nu(dx)[\varphi(x) - \varphi_1(x)]\, u(x, y),$$

and so

$$|\varphi(y) - \varphi_1(y)| \leqslant \int \nu(dx)\, |\varphi(x) - \varphi_1(x)|\, u(x, y). \tag{1.9}$$

But the integrals with respect to ν of the two members of (1.9) are equal by (1.3). Together with (1.9) this forces $\varphi - \varphi_1$ to be of a constant sign ν-a.e. Since $\int \varphi d\nu = 1 = \int \varphi_1 d\nu$, it follows that $\varphi = \varphi_1$ ν-a.e. and, therefore, $\pi = \pi_1$. This completes the proof.

Under the assumptions of [1], u is strictly positive, hence the function φ defined by (1.6) is also strictly positive. We shall assume this from now on in the general context. Put

$$u_\varphi(x, y) = \frac{u(x, y)}{\varphi(y)}. \tag{1.10}$$

Then (1.6) may be written as

$$\int \pi(dx)\, u_\varphi(x, y) = 1, \quad y \in K, \tag{1.11}$$

whereas, in view of (1.7), (1.3) may be written in the form

$$\int u_\varphi(x, y)\, \pi(dy) = 1, \quad x \in K. \tag{1.12}$$

We call u the *modified potential density* (relative to K) and record the next result as follows:

PROPOSITION 3. *There are a Borel function $\varphi > 0$ and a probability measure π on K satisfying* (1.11) *and* (1.12).

Let us tell the probabilistic origin of the symmetry exhibited in (1.11) and (1.12). Since π is an invariant probability measure for the Markovian kernel L, a *reverse kernel* in the most elementary (and classical) sense is given by

$$\hat{L}(y, dx) = \frac{\pi(dx) L(x, dy)}{\pi(dy)} = \pi(dx)\, u_\varphi(x, y).$$

Thus (1.11) states that $\hat{L}(y, \cdot)$ is a probability measure for each y. Now it is trivial that π is also an invariant measure for the Markovian kernel $\hat{L}$, namely, $\pi\hat{L} = \pi$. Written out this is just (1.12).

As an immediate application of (1.11), we mention the following extension of a familiar result in potential theory:

COROLLARY. *If μ is any measure such that*

$$\int_K u_\varphi(x, y)\, \mu(dy) \leqslant 1 \quad \textit{for all } x \in K,$$

then $\mu(K) \leqslant 1$.

Proof. Fubini's theorem yields

$$1 = \pi(K) \geqslant \int_K \pi(dx) \int_K u_\varphi(x, y)\mu(dy) \geqslant \int_K 1\mu(dy) = \mu(K).$$

Thus π fulfills the physicist's concept of the equilibrium distribution on K with respect to the modified potential u_φ, as well as the probabilist's one which is (1.12). We proceed to strengthen this analogy by considerations of energy.

2. Let λ be a signed finite measure on K; namely, $\lambda = \lambda^+ - \lambda^-$, where λ^+ and λ^- are measures on K such that $|\lambda| = \lambda^+ + \lambda^-$ is a finite measure. We denote this class of signed measures by $S(K)$. For $\lambda_1 \in S(K)$ and $\lambda_2 \in S(K)$, we define the *mutual energy* of λ_1 and λ_2 relative to u_φ by

$$\langle \lambda_1, \lambda_2 \rangle_\varphi = \iint \lambda_1(dx) u_\varphi(x, y) \lambda_2(dy) \tag{2.1}$$

with the stipulation that $\langle |\lambda_1|, |\lambda_2| \rangle_\varphi < \infty$; otherwise, the quantity in (2.1) is not defined and will not be written. For $\lambda \in S(K)$ we write

$$I_\varphi(\lambda) = \langle \lambda, \lambda \rangle_\varphi .$$

If $\lambda \in S(K)$ and $I_\varphi(|\lambda|) < \infty$, we call $I_\varphi(\lambda)$ the *energy* of λ and write $\lambda \in \mathscr{E}_\varphi$. The subclass of probability measures in $\mathscr{E}_\varphi$ will be denoted by $\mathscr{E}_\varphi^0$. Next, for $\lambda \in \mathscr{E}_\varphi$ we put

$$G_\varphi(\lambda) = I_\varphi(\lambda) - 2\lambda(1),$$

where, of course, $\lambda(1)$ may also be denoted by $\lambda(K)$. Then G_φ is the *Gauss quadratic*. It follows from (1.11) and (1.12) that

$$I_\varphi(\pi) = 1 \quad \text{and} \quad G_\varphi(\pi) = -1.$$

From here on the subscript φ will be omitted from these symbols except in u_φ, when there is no risk of confusion.

PROPOSITION 4. *If $\lambda \in \mathscr{E}$, then $\lambda + \pi \in \mathscr{E}$ and*

$$G(\lambda+\pi) = I(\lambda) + G(\pi) = I(\lambda) - 1. \tag{2.2}$$

Proof. We have

$$I(|\lambda+\pi|) \leqslant I(|\lambda|+\pi) = I(|\lambda|) + I(\pi) + \langle |\lambda|, \pi \rangle + \langle \pi, |\lambda| \rangle .$$

Now, by (1.12) and Fubini's theorem

$$\langle |\lambda|, \pi \rangle = \int |\lambda|(dx) \int u_\varphi(x, y) \pi(dy) = |\lambda|(1) < \infty;$$

similarly, by (1.11),

$$\langle \pi, |\lambda| \rangle = \int \left[\int \pi(dx) u_\varphi(x, y) \right] |\lambda|(dy) = |\lambda|(1) < \infty .$$

Thus the same calculations yield

$$I(\lambda+\pi) = I(\lambda)+I(\pi)+\langle\lambda,\pi\rangle+\langle\pi,\lambda\rangle = I(\lambda)+I(\pi)+2\lambda(1)$$

and (2.2) follows at once.

COROLLARY 1. *If λ is a probability measure on K such that $\lambda-\pi\in\mathscr{E}$, then $I(\lambda) = I(\pi)+I(\lambda-\pi)$.*

We say that I satisfies the *positivity principle* iff

$$I(\lambda) \geqslant 0 \quad \text{for every } \lambda\in\mathscr{E}_\varphi; \tag{2.3}$$

we say that I satisfies the *energy principle* iff (2.3) is true and, moreover, $I(\lambda) = 0$ implies $\lambda^+ \equiv \lambda^- \equiv 0$.

COROLLARY 2. *If I satisfies the positivity principle, then we have*

$$G(\pi) \leqslant G(\lambda) \quad \textit{for every } \lambda\in\mathscr{E}_\varphi, \tag{2.4}$$

$$I(\pi) \leqslant I(\lambda) \quad \textit{for every } \lambda\in\mathscr{E}^0_\varphi. \tag{2.5}$$

If I satisfies the energy principle, then π is the unique member of $\mathscr{E}_\varphi$ for which (2.4) *is true; and it is also the unique member of $\mathscr{E}^0_\varphi$ for which* (2.5) *is true.*

When u is symmetric, ν is invariant for L because

$$\int \nu(dx)u(x,y) = 1, \quad y\in K;$$

and we may normalize ν by putting

$$\pi = \frac{\nu}{\nu(1)}.$$

In this case $\varphi \equiv 1$, and if we use 1 as the subscript to indicate this case, we have

$$\langle\lambda_1,\lambda_2\rangle_1 = \frac{1}{\nu(1)}\iint \lambda_1(dx)u(x,y)\lambda_2(dy)$$

with the original potential density u. This is the classical situation and Corollary 2 contains the theorems on the minimization of energy by the equilibrium measure as given in the literature. For the Newtonian and Marcel Riesz potentials the energy principle is satisfied (see, e.g., [6]). Indeed, in the general symmetric case satisfying the energy principle it can be shown that the two ways of minimization in (2.4) and (2.5), respectively, are equivalent. We have not been able to trace the source of this fact, but the proof is standard.

A true symmetrization of u_φ may be considered by putting

$$\tilde{u}_\varphi(x,y) = \frac{1}{2}\left[\frac{u(x,y)}{\varphi(y)}+\frac{u(y,x)}{\varphi(x)}\right]. \tag{2.6}$$

Then (1.11) and (1.12) continue to hold when u_φ is replaced there by $\tilde{u}_\varphi$. Since $\tilde{u}_\varphi$ is symmetric, the well-known methods of potential theory apply. It will be seen in Section 4 that not only u_φ but also its transpose $u(y,x)/\varphi(x)$ is the potential density of a process, namely the dual process (see, e.g., [4]). But the operation of addition or averaging in (2.6) may not have a useful interpretation for probability theory, so we shall let it pass here.

3. Recall that the kernel L in (1.2) may be defined by

$$Lf(x) = \int u_\varphi(x, y) f(y) \pi(dy), \quad f \in \mathfrak{B}_+.$$

Its dual is defined by

$$\hat{L}f(y) = \int \pi(dx) f(x) u_\varphi(x, y), \quad f \in \mathfrak{B}_+.$$

Note that

$$\int (\hat{L}f) g d\pi = \int f(Lg) d\pi = \langle f \cdot \pi, g \cdot \pi \rangle,$$

where $f \cdot \pi$ is the measure $f(y)\pi(dy)$ and the notation above omits the subscript φ as before. Both L and $\hat{L}$ are contractions on $\mathscr{L}^p(\pi)$ for $1 \leqslant p \leqslant \infty$. To see this let $\int |f|^p d\pi < \infty$; then, since $|Lf|^p \leqslant L(|f|^p)$ and $\pi L = \pi$, we obtain

$$\int |Lf|^p d\pi \leqslant \int L|f|^p d\pi = \int |f|^p d\pi. \tag{3.1}$$

Note that it is sufficient that L be submarkovian and π subinvariant, $\pi L \leqslant \pi$, for (3.1) to hold. Similarly for $\hat{L}$.

Let λ be any probability measure on K, and put

$$g(y) = \int \lambda(dx) u_\varphi(x, y).$$

The measure λL is given by

$$\lambda L(dy) = \int \lambda(dx) L(x, dy) = g(y)\pi(dy).$$

An induction shows that, for $n \geqslant 1$,

$$\lambda L^n(dy) = \hat{L}^{n-1} g(y) \pi(dy).$$

We can now calculate, for $n \geqslant 1$ and $m \geqslant 1$,

$$\langle \lambda L^n, \lambda L^m \rangle = \iint \hat{L}^{n-1} g(x) \pi(dx) u_\varphi(x, y) \hat{L}^{m-1} g(y) \pi(dy) = \int \hat{L}^n g \cdot \hat{L}^{m-1} g d\pi.$$

This is also valid for $n = 0$ and $m \geqslant 1$ with, of course, $\hat{L}^0 g = g$. In particular,

$$\langle \lambda, \lambda L \rangle = \int g^2 d\pi, \tag{3.2}$$

$$\langle \lambda L, \lambda L \rangle = \int \hat{L}g \cdot g d\pi, \tag{3.3}$$

$$\langle \lambda L, \lambda L^2 \rangle = \int (\hat{L}g)^2 d\pi, \tag{3.4}$$

$$\langle \lambda L^2, \lambda L \rangle = \int L^2 g \cdot g d\pi = \int \hat{L}g \cdot Lg d\pi. \tag{3.5}$$

PROPOSITION 5. *The three quantities in* (3.3)-(3.5) *are all dominated by that in* (3.2). *In particular,* $\langle \lambda L^n, \lambda L^{n+1}\rangle$ *decreases as n increases.*

Proof. We have

$$\int \hat{L}g\cdot g d\pi \leqslant \left[\int (\hat{L}g)^2 d\pi \int g^2 d\pi\right]^{1/2} \leqslant \int g^2 d\pi$$

by the Cauchy-Schwarz inequality followed by (3.1) for $\hat{L}$ with $p = 2$. The rest is similar. Having shown that $\langle \lambda L, \lambda L^2\rangle \leqslant \langle \lambda, \lambda L\rangle$ for any λ, an iteration establishes the last assertion in Proposition 5.

In the symmetry case, it is trivial that the positivity principle implies the Cauchy-Schwarz inequality for mutual energy as follows:

$$(3.6) \qquad \langle \lambda_1, \lambda_2\rangle^2 \leqslant \langle \lambda_1, \lambda_1\rangle\langle \lambda_2, \lambda_2\rangle.$$

In general it is not clear when this is valid.

COROLLARY. *If I satisfies the Cauchy-Schwarz inequality* (3.6), *then*

(*) $\langle \lambda L^n, \lambda L^n\rangle$ *decreases as n increases for* $n \geqslant 0$.

Proof. We have $\langle \lambda, \lambda L\rangle^2 \leqslant \langle \lambda, \lambda\rangle\langle \lambda L, \lambda L\rangle$. Together with $\langle \lambda L, \lambda L\rangle \leqslant \langle \lambda, \lambda L\rangle$, as given in Proposition 5, we obtain $\langle \lambda L, \lambda L\rangle \leqslant \langle \lambda, \lambda\rangle$. This implies (*) upon iteration.

For a symmetric u (i.e., $\varphi \equiv 1$), Proposition 5 is due to J. R. Baxter and the Corollary answers a conjecture by J. B. Walsh. Under stronger conditions on u (see, e.g., [5]), so that $\lambda L^n \to \pi$ as $n \to \infty$, it appears that (*) should imply the limiting relation $\langle \pi, \pi\rangle \leqslant \langle \lambda, \lambda\rangle$. But this implication is actually false in general because u is not bounded. On the other hand, the last written inequality has been proved directly in Corollary 2 to Proposition 4.

4. It is common knowledge that if u is the potential density of the Markov process $\{X_t, t \geqslant 0\}$, then the modified u_φ defined in (1.11) is that of another Markov process $\{Y_t, t \geqslant 0\}$ obtained from X by a random time change. More precisely, let

$$\tau(t) = \int_0^t \frac{1}{\varphi(X_s)}\, ds$$

defined for each sample path. Under the assumptions of [1], $\{X_t\}$ is a transient Hunt process and φ is lower semi-continuous and $\varphi > 0$. Thus, for each t, $\varphi(X_s)$ is bounded away from zero for $0 \leqslant s \leqslant t$. It follows that $\tau(t) < \infty$ for each t, almost surely. Thus $\tau(t)$ is continuous non-decreasing in t, and so has a right continuous, strictly increasing inverse τ^{-1}. Define Y by

$$Y_t = X_{\tau^{-1}(t)}, \quad t \geqslant 0.$$

Then Y is a right continuous strong Markov process in $[0, \infty)$. For every positive measurable f we have

$$\int_0^\infty f(Y_t)\,dt = \int_0^\infty f(X_t)\frac{1}{\varphi(X_t)}\,dt.$$

Hence, if we write U_X and U_Y for the potential kernel of X and Y, respectively, we obtain

$$U_Y f = \int u_\varphi(x, y) f(y) m(dy).$$

Thus Y has the potential density $u_\varphi(x, y)m(dy)$.

We now discuss the positivity principle for I_φ in the context of additive functionals. Consider a natural increasing additive functional $\{A_t, t > 0\}$. Its potential $U_A 1$ has a representation $U\mu_A$, by Theorem 2, Corollary 2, of [3]. This measure can be shown to be the Revuz measure associated with A (see [7]). For positive measurable f we write $f_K = f 1_K$; then we have

$$\text{(4.1)}\qquad \mathrm{E}^x\{\int_0^\infty f(X_{s-})1_K(X_{s-})\,dA_s\} = U_A f_K(x) = \int_K u(x, y) f(y)\mu_A(dy).$$

Taking $f = 1/\varphi$, where φ is the function in (1.7), and the difference of two such functionals, we obtain in obvious notation the formula

$$\mathrm{E}^x\left\{\int_0^\infty \frac{1_K(X_{s-})}{\varphi(X_{s-})}\,d(A_s^+ - A_s^-)\right\} = \int_K u_\varphi(x, y)\mu_A(dy),$$

where $A = A^+ - A^-$. Putting

$$dB_s = \frac{1_K(X_{s-})}{\varphi(X_{s-})}\,dA_s,$$

we have

$$h(x) \overset{\text{def}}{=} \mathrm{E}^x\{B_\infty\} = \int_K u_\varphi(x, y)\mu_A(dy).$$

It follows by a familiar calculation that

$$\mathrm{E}^x\{B_\infty^2\} = \mathrm{E}^x\{\int_0^\infty [\int_{[t,\infty)} dB_s + \int_{(t,\infty)} dB_s]\,dB_t\} = \mathrm{E}^x\{\int_0^\infty [h(X_t)_- + h(X_t)]\,dB_t\},$$

where

$$h(X_t)_- = \lim_{s\uparrow\uparrow t} h(X_s).$$

Since h is the difference of two excessive functions, the limit exists and is equal to $h(X_t)$ except for a countable set of t (depending on the path). Hence, if B is a continuous additive functional, then the above is equal to

$$2\mathrm{E}^x\{\int_0^\infty h(X_t)\,dB_t\} = 2U_B h(x) = 2\int_K u_\varphi(x, y) h(y)\mu_A(dy).$$

Thus we obtain

$$E^x\{B_\infty^2\} = 2\int_K u_\varphi(x, y)\int_K u_\varphi(y, z)\mu_A(dz)\mu_A(dy),$$

which shows that the iterated integral above has a value greater than or equal to 0. Now, if we integrate it with respect to π and use (1.12), the result is

$$I_\varphi(\mu_A) = \int_K\int_K \mu_A(dy)u_\varphi(y, z)\mu_A(dz) \geqslant 0.$$

We have proved the following

PROPOSITION 6. *For every measure μ_A associated with a continuous additive functional A of the process X as in* (4.1), *we have $I_\varphi(\mu_A|_K) \geqslant 0$, where $\mu_A|_K$ is the restriction of μ_A to the compact K.*

The minimization results of (2.4) and (2.5) therefore hold true for this class of measures. The question whether $I_\varphi(\mu_A) = 0$ implies $\mu_A \equiv 0$ seems more difficult and remains to be investigated.

Added in proof. S. Orey informed us of the following complement to Proposition 2:

If an invariant probability measure π exists as in (1.4), *then the Markov chain associated with the kernel L is ν-recurrent and, for each x,* $\lim_n L^{(n)}(x, A) = \pi(A)$.

The proof is simple.

REFERENCES

[1] K. L. Chung, *Probabilistic approach in potential theory to the equilibrium problem*, Ann. Inst. Fourier (Grenoble) 23 (1973), p. 313-322.

[2] — *Remarks on equilibrium potential and energy*, ibidem 25 (1975), p. 131-138.

[3] — and M. Rao, *A new setting for potential theory*, Part I (to appear).

[4] — *On existence of a dual process*, Matematisk Institut, Aarhus University, Preprint Series No. 25 (1978).

[5] J. L. Doob, *Stochastic processes*, J. Wiley, 1953.

[6] M. Rao, *Brownian motion and classical potential theory*, Matematisk Institut, Aarhus University Lecture Notes Series No. 47 (1977).

[7] — *A note on Revuz measure*, Matematisk Institut, Aarhus Universitet, Preprint Series No. 19 (1978-1979).

[8] D. Revuz, *Mesures associées aux fonctionelles additives de Markov, I*, Trans. Amer. Math. Soc. 148 (1970), p. 501-531.

Department of Mathematics
Stanford University
Stanford, California 94305, U.S.A.

Received on 12. 11. 1979

STUDIA MATHEMATICA, T. LXVIII. (1980)

Kac functional and Schrödinger equation

by

KAI LAI CHUNG* and S. R. S. VARADHAN (Stanford, Calif.)

1. Introduction. Let $X = \{x(t),\ t \geqslant 0\}$ be a strong Markov process with continuous paths on $\boldsymbol{R} = (-\infty, +\infty)$. Such a process is often called a *diffusion*. For each real b, we define the hitting time τ_b as follows:

$$\tau_b = \inf\{t > 0 \mid x(t) = b\}. \tag{1}$$

Let P_a^1 and E_a denote as usual the basic probability and expectation associated with paths starting from a. It is assumed that for every a and b, we have

$$P_a\{\tau_b < \infty\} = 1. \tag{2}$$

Now let q be a bounded Borel measurable function on $\boldsymbol{R}$, and write for brevity

$$e(t) = \exp \int_0^t q\big(x(s)\big)\,ds. \tag{3}$$

This is a multiplicative functional introduced by M. Kac in [3]. In this paper we study the quantity

$$u(a, b) = E_a\{e(\tau_b)\}. \tag{4}$$

Since q is bounded below, (2) implies that $u(a, b) > 0$ for every a and b, but it may be equal to $+\infty$. A fundamental property of u is given by

$$u(a, b)u(b, c) = u(a, c), \tag{5}$$

valid for $a < b < c$, or $a > b > c$. This is a consequence of the strong Markov property (SMP).

2. Probabilistic investigation. We begin by defining two abscissas of finiteness, one for each direction.

$$\begin{aligned} \beta &= \inf\{b \in \boldsymbol{R} \mid \exists a < b:\ u(a, b) = \infty\} \\ &= \sup\{b \in \boldsymbol{R} \mid \forall a < b:\ u(a, b) < \infty\}; \\ \alpha &= \sup\{a \in \boldsymbol{R} \mid \exists b > a:\ u(b, a) = \infty\} \\ &= \inf\{a \in \boldsymbol{R} \mid \forall b > a:\ u(b, a) < \infty\}. \end{aligned} \tag{6}$$

* Research supported in part by NSF Grant MCS77-01319.

It is possible, e.g., that $\beta = -\infty$ or $+\infty$. The first case occurs when X is the standard Brownian motion, and $q(x) \equiv 1$; for then, $u(a, b) \geqslant E_a(\tau_b) = \infty$, for any $a \neq b$.

LEMMA 1. *We have*

$$\begin{aligned}\beta &= \inf\{b \in \boldsymbol{R} \mid \forall a < b\colon\ u(a, b) = \infty\}\\ &= \sup\{b \in \boldsymbol{R} \mid \exists a < b\colon\ u(a, b) < \infty\};\\ \alpha &= \sup\{a \in \boldsymbol{R} \mid \forall b > a\colon\ u(b, a) = \infty\}\\ &= \inf\{a \in \boldsymbol{R} \mid \exists b > a\colon\ u(b, a) < \infty\}.\end{aligned}$$

Proof. It is sufficient to prove the first equation above for β, because the second is trivially equivalent to it, and the equations for α follow by similar arguments. Suppose $u(a, b) = \infty$; then for $x < a < b$ we have $u(x, b) = \infty$ by (5). For $a < x < b$ we have by SMP,

$$u(x, b) \geqslant E_x\{e(\tau_a);\ \tau_a < \tau_b\}\, u(a, b) = \infty$$

since $P_x\{\tau_a < \tau_b\} > 0$ in consequence of (2).

The next lemma is a martingale argument. Let $\mathfrak{J}_t$ be the σ-field generated by $\{x_s,\ 0 \leqslant s \leqslant t\}$ and all null sets, so that $\mathfrak{J}_{t+} = \mathfrak{J}_t$ for $t \geqslant 0$; and for any optional τ let $\mathfrak{J}_{\tau+}$, $\mathfrak{J}_\tau$, $\mathfrak{J}_{\tau-}$ have the usual meanings.

LEMMA 2. *If* $a < b < \beta$, *then*

$$\lim_{a \uparrow b} u(a, b) = 1; \tag{7}$$

$$\lim_{b \downarrow a} u(a, b) = 1. \tag{8}$$

Proof. Let $a < b_n \uparrow b$ and consider

$$E_a\{e(\tau_b) \mid \mathfrak{J}(\tau_{b_n})\}, \quad n \geqslant 1. \tag{9}$$

Since $b < \beta$, $u(a, b) < \infty$ and the sequence in (9) forms a martingale. As $n \uparrow \infty$, $\tau_{b_n} \uparrow \tau_b$ a.s. and $\mathfrak{J}(\tau_{b_n}) \uparrow \mathfrak{J}(\tau_b^-)$. Since $e(\tau_b) \in \mathfrak{J}(\tau_b^-)$, the limit of the martingale is a.s. equal to $e(\tau_b)$. On the other hand, the conditional probability in (9) is also equal to

$$E_a\left\{e(\tau_b)\exp\left(\int_{\tau_{b_n}}^{\tau_b} q\big(x(s)\big)\,ds\right) \Big| \mathfrak{J}(\tau_{b_n})\right\} = e(\tau_{b_n})\, u(b_n, b).$$

As $n \uparrow \infty$, this must then converge to $e(\tau_b)$ a.s.; since $e(\tau_{b_n})$ converges to $e(\tau_b)$ a.s., we conclude that $u(b_n, b) \to 1$. This establishes (7).

Now let $\beta > b > a_n \downarrow a$, and consider

$$E_a\{e(\tau_b) \mid \mathfrak{J}(\tau_{a_n})\}, \quad n \geqslant 1. \tag{10}$$

This is again a martingale. Although $a \to \tau_a$ is a.s. left continuous, not right continuous, for each fixed a we do have $\tau_{a_n} \downarrow \tau_a$ and $\mathfrak{J}(\tau_{a_n}) \downarrow \mathfrak{J}(\tau_a)$.

Hence we obtain as before $u(a_n, b) \to u(a, b)$ and consequently

$$u(a, a_n) = \frac{u(a, b)}{u(a_n, b)} \to 1.$$

This establishes (8).

The next result illustrates the basic probabilistic method.

THEOREM 1. *The following three propositions are equivalent*:

(i) $\beta = +\infty$;

(ii) $\alpha = -\infty$;

(iii) *For every a and b, we have*

$$u(a, b)u(b, a) \leqslant 1. \tag{11}$$

Proof. Suppose $x(0) = b$ and let $a < b < c$. If (i) is true, then $u(b, c) < \infty$ for every $c > b$. Define a sequence of successive hitting times T_n as follows (where θ denotes the usual shift operator):

$$S = \begin{cases} \tau_a & \text{if} \quad \tau_a < \tau_c, \\ \infty & \text{if} \quad \tau_c < \tau_a; \end{cases}$$

$$T_0 = 0, \quad T_1 = S, \tag{12}$$

$$T_{2n} = T_{2n-1} + \tau_b \circ \theta_{T_{2n-1}}, \quad T_{2n+1} = T_{2n} + S \circ \theta_{T_{2n}},$$

for $n \geqslant 1$. Define also

$$N = \min\{n \geqslant 0 \mid T_{2n+1} = \infty\}. \tag{13}$$

It follows from $P_b\{\tau_c < \infty\} = 1$ that $0 \leqslant N < \infty$ a.s. For $n \geqslant 0$, we have

$$E_b\{e(\tau_c) \colon N = n\} = E_b\Big\{\exp\Big(\sum_{k=0}^{2n} \int_{T_k}^{T_{k+1}} q(x(s))\,ds\Big)\Big\} \tag{14}$$

$$= E_b\{e(\tau_a);\ \tau_a < \tau_c\}^n\, E_a\{e(\tau_b)\}^n E_b\{e(\tau_c);\ \tau_c < \tau_a\}.$$

Since the sum of the first term in (14) over $n \geqslant 0$ is equal to $u(b, c) < \infty$, the sum of the last term in (14) must converge. Thus we have

$$E_b\{e(\tau_a);\ \tau_a < \tau_c\} u(a, b) < 1. \tag{15}$$

Letting $c \to \infty$ we obtain (11). Hence $u(b, a) < \infty$ for every $a < b$ and so (ii) is true. Exactly the same argument shows that (ii) implies (iii) and so also (i).

We are indebted to R. Durrett for ridding the next lemma of a superfluous condition.

LEMMA 3. *Given any $a \in \boldsymbol{R}$ and $Q > 0$, there exists an $\varepsilon = \varepsilon(a, Q)$ such that*

$$E_a\{e^{Q\sigma_\varepsilon}\} < \infty \tag{16}$$

where $\sigma_\varepsilon = \inf\{t > 0 \mid x(t) \notin (a-\varepsilon, a+\varepsilon)\}$.

Proof. Since X is strong Markov and has continuous paths, there is no "stable" point. This implies $P_a\{\sigma_\varepsilon \geqslant 1\} \to 0$ as $\varepsilon \to 0$ and so there exists ε such that

$$P_a\{\sigma_\varepsilon \geqslant 1\} < e^{-(Q+1)}. \tag{17}$$

Now σ_ε is a terminal time, so $x \to P_x\{\sigma_\varepsilon \geqslant 1\}$ is an excessive function for the process X killed at σ_ε. Hence by standard theory it is finely continuous. By for a diffusion under hypothesis (2) it is clear that fine topology coincides with the Euclidean. Thus $x \to P_x\{\sigma_\varepsilon \geqslant 1\}$ is in fact continuous. (1) It now follows that we have, further decreasing ε if necessary:

$$\sup_{|x-a|<\varepsilon} P_x\{\sigma_\varepsilon \geqslant 1\} < e^{-(Q+1)}. \tag{18}$$

A familiar inductive argument then yields for all $n \geqslant 1$

$$P_a\{\sigma_\varepsilon \geqslant n\} < e^{-n(Q+1)} \tag{19}$$

and (16) follows.

LEMMA 4. *For any* $a < \beta$ *we have*

$$u(a, \beta) = \infty; \tag{20}$$

for any $b > \alpha$ *we have* $u(b, \alpha) = \infty$.

Proof. We will prove that if $u(a, b) < \infty$, then there exists $c > b$ such that $u(b, c) < \infty$. This implies (20) by Lemma 1, and the second assertion is proved similarly.

Let $Q = \|q\|$. Given b we choose a and b so that $a < b < d$ and

$$E_b\{e^{Q(\tau_a \wedge \tau_d)}\} < \infty. \tag{21}$$

This is possible by Lemma 3. Now let $b < c < d$; then as $c \downarrow b$ we have

$$E_b\{e(\tau_a);\ \tau_a < \tau_c\} \leqslant E_b\{e^{Q(\tau_a \wedge \tau_d)};\ \tau_a < \tau_c\} \to 0 \tag{22}$$

because $P_b\{\tau_a < \tau_c\} \to 0$. Hence there exists c such that

$$E_b\{e(\tau_a);\ \tau_a < \tau_c\} < \frac{1}{u(a, b)}. \tag{23}$$

This is just (15) above, and so reversing the argument there, we conclude that the sum of the first term in (14) over $n \geqslant 0$ must converge. Thus $u(b, c) < \infty$, as was to be shown.

To sum up:

THEOREM 2. *The function* $(a, b) \to u(a, b)$ *is continuous in the region* $a \leqslant b < \beta$ *and in the region* $\alpha < b \leqslant a$. *Furthermore, extended continuity*

---

(1) This fact can also be proved in an elementary way.

holds in $a \leqslant b \leqslant \beta$ *and* $\alpha \leqslant b \leqslant a$, *except at* (β, β) *when* $\beta < \infty$, *and at* (α, α) *when* $\alpha > -\infty$.

Proof. To see that there is continuity in the extended sense at (a, β), where $a < \beta$, let $a < b_n \uparrow \beta$. Then we have by Fatou's lemma

$$\varliminf_{n\to\infty} u(a, b_n) \geqslant E_a\{\lim_{n\to\infty} e(\tau_{b_n})\} = E_a\{e(\tau_\beta)\} = u(a, \beta) = \infty.$$

If $\beta < \infty$, then $u(\beta, \beta) = 1$ by definition, but $u(a, \beta) = \infty$ for all $a < \beta$; hence u is not continuous at (β, β). The case for α is similar.

3. The Schrödinger equation. From now on the process X will be the standard Brownian motion on $\boldsymbol{R}$ and q will be bounded and continuous in $\boldsymbol{R}$. The Schrödinger equation

$$\frac{1}{2}\varphi''(x) + q(x)\varphi(x) = 0, \quad x \in \boldsymbol{R} \tag{24}$$

will be referred to as "the equation", and any of its solutions "a solution". The fundamental existence and uniqueness theorem for linear differential equations (with a Lipshitz condition) is applicable and guarantees a unique solution for given initial values $\varphi(a)$ and $\varphi'(a)$ for any $a \in \boldsymbol{R}$. A fortiori, it guarantees the unique extension of any solution given in a non-empty interval to a solution in $\boldsymbol{R}$.

LEMMA 5. *Let* φ *be a solution in* $[a, b]$ *with* $\varphi(a) = \varphi(b) = 0$; *then* $b - a \geqslant (2Q)^{-1/2}$, *where* $Q = \|q\|$. *Let* $0 < b - a < (2Q)^{-1/2}$, $A > 0$, $B > 0$; *then there is a unique solution* φ *satisfying*

$$\varphi(a) = A, \quad \varphi(b) = B, \tag{25}$$

and $\varphi > 0$ *in* $[a, b]$.

Proof. The first assertion is known as de la Vallée Poussin's theorem. The second assertion then follows from a well-known criterion for the solvability of the equation with given boundary conditions. For both results see [5], pp. 91–92. The solution cannot vanish more than once by the first assertion. Nor can it vanish just once for then it must assume its minimum there and so Φ' must also vanish which is impossible by the uniqueness theorem.

LEMMA 6. *Let* Φ *be a solution in* $\boldsymbol{R}$. *For* $-\infty < a < b < \infty$ *define*

$$\tau = \inf\{t \geqslant 0 \mid x(t) \notin [a, b])\} \tag{26}$$

and

$$M(t) = \Phi\big(x(t)\big) \exp \int_0^t q\big(x(s)\big)\, ds. \tag{27}$$

Then $\{M(t \wedge \tau), \mathfrak{J}(t), t \geqslant 0\}$ *is a martingale.*

Proof. Using Ito's formula (see [2]), we have

$$(28)\qquad dM(t) = \exp\int_0^t q(x(s))\,ds \times \\ \times\Big[\Phi'(x(t))\,dx(t) + \big(\tfrac{1}{2}\Phi''(x(t)) + q(x(t))\,\Phi(x(t))\big)\,dt\Big],$$

namely, for $t \geqslant 0$:

$$(29)\qquad M(t)-M(0) = \int_0^t \Big[\exp\int_0^s q(x(r))\,dr\Big]\Phi'(x(s))\,dx(s) + \\ + \int_0^t \Big[\exp\int_0^s q(x(r))\,dr\Big]\big[\tfrac{1}{2}\Phi''+q\Phi\big](x(s))\,ds.$$

If we substitute $t\wedge\tau$ for t in the above, the second term on the right vanishes because Φ is a solution. The first term is then of the form $\int_0^t f(s)\,dx(s)$ where

$$f(s) = \chi(s)\Big[\exp\int_0^s q(x(r))\,dr\Big]\,\Phi'(x(s)),$$

$$(30)\qquad \chi(s, w) = \begin{cases} 1 & \text{if} \quad s < \tau(w), \\ 0 & \text{if} \quad s \geqslant \tau(w). \end{cases}$$

Clearly, $f(s, w)$ is progressively measurable, being right continuous, and

$$(31)\qquad E_x\Big\{\int_0^t f(x)^2\,ds\Big\} < \infty, \qquad x \in \boldsymbol{R},$$

because φ' as well as φ is bounded in $[a, b]$. Thus the first term on the right side of (29) is an Ito integral, hence a martingale.

THEOREM 3. *Suppose $u(x, b) < \infty$ for some, hence all, $x < b$. Then $u(\cdot, b)$ is a solution in $(-\infty, b)$.*

Proof. Let $x_1 < x < x_2 < b$ where $x_2 - x_1 < (2Q)^{-1/2}$. Then by Lemma 5, there is a solution Φ satisfying

$$(32)\qquad \Phi(x_i) = u(x_i, b), \qquad i = 1, 2.$$

Let

$$\sigma = \inf\{t \geqslant 0 \mid x(t) \notin [x_1, x_2]\}$$

and M be as in (27). Then $M(t\wedge\sigma)$ is a martingale by Lemma 6. Hence for $t \geqslant 0$ we have

$$(33)\qquad \Phi(x) = E_x\{M(0)\} = E_x\{M(t\wedge\sigma)\} \\ = E_x\{M(\sigma);\ \sigma < t\} + E_x\{M(t);\ \sigma \geqslant t\}.$$

By Lemma 5, we can choose $x_2 - x_1$ so small that

(34) $$E\,\{e^{Q\sigma}\} < \infty.$$

Then

$$E_x\{M(t);\ \sigma \geqslant t\} \sup_{x_1 \leqslant x \leqslant x_2} |\Phi| \int_{\{\sigma \geqslant t\}} e^{Q\sigma} dP_x$$

converges to zero as $t \to \infty$ by (34), since $\sigma < \infty$ almost surely. Therefore, if we let $t \to \infty$ in (33) and use (32) we obtain

(35) $$\Phi(x) = E_x\{M(\sigma)\} = E_x\{u(x(\sigma),\, b)\, e(\sigma)\}.$$

On the other hand, by the strong Markov property applied at σ $(< \tau_b)$, we see that

(36) $$\begin{aligned} u(x, b) &= E_x\{e(\sigma) E_{x(\sigma)}[e(\tau_b)]\} \\ &= E_x\{e(\sigma) u(x(\sigma), b)\}. \end{aligned}$$

Comparison of (35) with (36) yields $u(x, b) = \Phi(x)$. This being true for each $x < b$, Theorem 3 is established. See the Remarks at the end of the paper.

THEOREM 4. *Let Φ be any solution such that $\Phi(x) > 0$ for $x \in (-\infty, b]$. Then $b < \beta$, and we have*

(37) $$u(x, b) \leqslant \frac{\Phi(x)}{\Phi(b)}, \qquad -\infty < x \leqslant b.$$

In other words, $u(\cdot, b)$ is the minimal positive solution in $(-\infty, b)$ with $\lim_{x \uparrow b} u(x, b) = 1$.

Proof. Consider the M in (27) but write $\tau_{[a,b]}$ for the τ in (26). Then for each $x \in \boldsymbol{R}$ and $t \geqslant 0$, we have the martingale relation

(38) $$\Phi(x) = E_x\{M(0)\} = E_x\{M(t \wedge \tau_{[a,b]})\}.$$

If we keep b fixed but let $a \to -\infty$, then under P_x for $x \in (-\infty, b)$ we have $\tau_{[a,b]} \uparrow \tau_b$ and $M(t \wedge \tau_{[a,b]}) \to M(t \wedge \tau_b)$ by continuity. Since $M(t \wedge \tau_{[a,b]}) > 0$ because $\Phi > 0$ in $(-\infty, b]$, it follows from (38) by Fatou's lemma that

(39) $$\Phi(x) \geqslant E_x\{M(t \wedge \tau_b)\}, \qquad -\infty < x < b.$$

Letting $t \to \infty$, we obtain by another application of Fatou's lemma

(40) $$\Phi(x) \geqslant E_x\{M(\tau_b)\} = \Phi(b) u(x, b).$$

Thus (37) is true for $x < b$; for $x = b$ it reduces to a trivial equation.

As a corollary to Theorem 4, we can relate the two abscissas β and α to the solutions $u(\cdot, b)$. Suppose $\beta > -\infty$. For each $b < \beta$, $u(\cdot, b)$ is a solution in $(-\infty, b)$ by Theorem 3. By the fundamental existence and

uniqueness theorem for linear differential equations, there is a solution Φ_b in $\boldsymbol{R}$ satisfying

$$\Phi_b(x) = u(x, b), \quad -\infty < x < b. \tag{41}$$

It follows from (7) that $\Phi_b(b) = 1$. Furthermore, the uniqueness theorem implies that for $b < c < \beta$ we have

$$\Phi_c(x) = \Phi_b(x) u(b, c), \quad x \in \boldsymbol{R}. \tag{42}$$

Thus the family of solutions $\{\Phi_b,\ b < \beta\}$ are linearly dependent.

Similarly if $\alpha < +\infty$, then there is a family of linearly dependent solutions $\{{}_a\Phi,\ a > \alpha\}$ satisfying

$${}_a\Phi(x) = u(x, a), \quad a < x < +\infty. \tag{43}$$

COROLLARY TO THEOREM 4. *For each $b < \beta$, β is the smallest root of Φ_b. For each $a > \alpha$, α is the largest root of ${}_a\Phi$.*

Proof. Fix $b < \beta$ and denote the smallest root of Φ_b by z. Since $\Phi_b > 0$ in $(-\infty, z)$, we must have $z \leqslant \beta$ by Theorem 4. On the other hand, for $b < c < \beta$ we have $\Phi_b(x) = \varphi_c(x)/u(b, c) > 0$ for $x \in (-\infty, c)$ by (42).

Hence $\beta \leqslant z$ and so $\beta = z$ as asserted. The assertion about α is proved similarly.

THEOREM 5. *The following propositions are all equivalent:*

(i) *There is a solution which is positive in $\boldsymbol{R}$;*

(ii) $\beta = +\infty$;

(iii) $\alpha = -\infty$;

(iv) *for every pair of real numbers a and b we have*

$$u(a, b) u(b, a) \leqslant 1; \tag{44}$$

(v) *for a pair of real numbers a and b we have* (44).

Proof. The equivalence of (ii), (iii) and (iv) has already been proved for any diffusion in Theorem 1.

Let Φ be a positive solution in $\boldsymbol{R}$. Then Theorem 4 applies for every b in $\boldsymbol{R}$ and yields $\beta = +\infty$. By symmetry $\alpha = -\infty$.

If $\beta = +\infty$, then for any $b \in \boldsymbol{R}$, Φ_b has no root by Corollary to Theorem 4; hence it is a positive solution in $\boldsymbol{R}$. Similarly if $\alpha = -\infty$, then for any $a \in \boldsymbol{R}$, ${}_a\Phi$ is a positive solution in $\boldsymbol{R}$. We have thus proved the equivalence of (i) with (ii), (iii), and (iv).

It remains to prove that (v) implies (i). Let a and b satisfy (44) and $a < b$. Then

$$\alpha < a < b < \beta \tag{45}$$

by the definition of α and β and Lemma 4. Bisecting the interval $[a, b]$

we deduce from (44) and (5) that $u(a_1, b_1)u(b_1, a_1) \leqslant 1$ for some a_1 and b_1 where $a \leqslant a_1 \leqslant b_1 \leqslant b$ and $b_1 - a_1 = (b-a)/2$. Continuing this process in the grand tradition of Bolzano–Weierstrass we see that there exist a_n and b_n such that $a_n \uparrow c$, $b_n \downarrow c$ where $c \in [a, b]$, and $u(a_n, b_n)u(b_n, a_n) \leqslant 1$. Since $a \leqslant a_n \leqslant b_n \leqslant b$ we may use (41), and (43) and (5) to obtain

$$\Phi_b(a_n)\,{}_a\Phi(b_n) \leqslant \Phi_b(b_n)\,{}_a\Phi(a_n). \tag{46}$$

Writing, e.g., $\Phi_b(b_n) = \Phi_b(a_n) + \Phi_b'(c_n)(b_n - a_n)$ where $a_n \leqslant c_n \leqslant b_n$ by the mean value theorem, substituting into (46) and letting $n \to \infty$, we obtain

$$W(c) = {}_a\Phi(c)\Phi_b'(c) - \Phi_b(c)\,{}_a\Phi'(c) \geqslant 0, \tag{47}$$

where W is the Wronskian of ${}_a\Phi$ and Φ_b. Since ${}_a\Phi$ and Φ_b are both solutions of the equation, it is an elementary fact that W is a constant. Now suppose for the sake of contradiction that Φ_b has a root; then by Corollary to Theorem 4 its smallest root is β. Since $\Phi_b > 0$ in $(-\infty, \beta)$, it is obvious that $\Phi_b'(\beta) < 0$; since $\beta > a$, we have ${}_a\Phi(\beta) > 0$ by the definition of ${}_a\Phi$. Thus

$$W(\beta) = {}_a\Phi(\beta)\Phi_b'(\beta) < 0, \tag{48}$$

which contradicts (47). Therefore, φ_b is a positive solution in $\boldsymbol{R}$ and (i) is proved.

Acknowledgement. Under the assumption (49) below, van Moerbeke proved that condition (v) implies that the solution w_1 in (50) has no zero. The proof above is modelled after his.

4. A particular case. In the analytical study of the Schrödinger equation (24), the following condition on q is often assumed:

$$\int_{-\infty}^{\infty} |xq(x)|\,dx < \infty. \tag{49}$$

It is known (see e.g. [1], p. 284) that there are then two solutions w_1 and w_2 such that

$$\lim_{x \to -\infty} w_1(x) = 1, \quad \lim_{x \to +\infty} w_2(x) = 1. \tag{50}$$

Any solution v which tends to a finite limit as $x \to -\infty$ $[+\infty]$ must be a constant multiple of w_1 $[w_2]$. For v' must tend to zero at $-\infty$ $[+\infty]$, so that Wronskian of v and w_1 $[w_2]$ must vanish.

The probabilistic counterpart of the result above is given below.

THEOREM 6. *Under the assumption* (49) *we have for any* $b < \beta$,

$$\lim_{x \to -\infty} u(x, b) < \infty \tag{51}$$

and for any $a > \alpha$,

(52) $$\lim_{x\to+\infty} u(x, a) < \infty.$$

Proof. We use the trivial inequality, for $x \leqslant b$,

(53) $$|u(x,b)-1| \leqslant \sum_{n=1}^{\infty} \frac{1}{n!} E_x\Big\{\Big[\int_0^{\tau_b} |q|\big(x(s)\big)\,ds\Big]^n\Big\} = \sum_{n=1}^{\infty} M^{(n)}(x, b),$$

where $(t_0 = 0)$

(54) $$M^{(n)}(x, b) = E_x\Big\{\prod_{j=1}^{n} \int_{t_{j-1}}^{\tau_b} |q|\big(x(t_j)\big)\,dt_j\Big\}.$$

Put also

$$M^{(n)}(b) = \sup_{x\leqslant b} M^{(n)}(x, b),$$

then $M^{(n)}(b)$ is nondecreasing in b. Using the Markov property of x at t_{n-1} in (54), we obtain

(55) $$M^{(n)}(x, b) \leqslant M^{(n-1)}(x, b)\, M^{(1)}(b) \leqslant [M^{(1)}(b)]^n.$$

Let τ be as in (26). Then it is part of the standard theory of Brownian motion that for $x \in [a, b]$ we have

(56) $$E_x\Big\{\int_0^{\tau} |q|\big(x(t)\big)\,dt\Big\} = \int_a^b G(x, y)|q|(y)\,dy,$$

where G is the Green's function for $[a, b]$, specifically (see, e.g. [2]),

$$G(x, y) = \begin{cases} \dfrac{(x-a)(b-y)}{b-a} & \text{if} \quad a \leqslant x \leqslant y \leqslant b; \\[2ex] \dfrac{(y-a)(b-x)}{b-a} & \text{if} \quad a \leqslant y \leqslant x \leqslant b. \end{cases}$$

Letting $a \to -\infty$ and using (49), we have by dominated convergence

(57) $$M^{(1)}(x, b) = E_x\Big\{\int_0^{\tau_b} |q|\big(x(t)\big)\,dt\Big\} = \int_{-\infty}^{b} [(b-x)\wedge(b-y)]\,|q|(y)\,dy.$$

Hence

(58) $$M^{(1)}(b) \leqslant \int_{-\infty}^{b} (b-y)\,|q|(y)\,dy$$

which tends to zero as $b \to -\infty$ by (49). Choose b_0 so that

(59) $$M^{(1)}(b_0) = \eta < 1.$$

Then we have by (53), (55) and (59) for $b < b_0$:

$$\sup_{x\leqslant b}|u(x,\ b)-1|\leqslant \frac{M^{(1)}(b)}{1-\eta},$$

and consequently the left member above tends to zero as $b\to-\infty$. This and the fundamental relation (5) for u implies (51) for $b=b_0$, hence also for all $b<\beta$ by the meaning of β (Lemma 1). This proves (51), and (52) is entirely similar.

COROLLARY. *We have the following bounds for β and α:*

$$\beta \geqslant \sup\Big\{b\in \boldsymbol{R}\,\Big|\ \int_{-\infty}^{b}(b-y)|q|(y)\,dy\ \leqslant 1\Big\},$$

$$\alpha\leqslant \inf\Big\{a\in \boldsymbol{R}\,\Big|\ \int_{a}^{\infty}(y-a)|q|(y)\,dy\leqslant 1\Big\}. \tag{60}$$

5. Remarks. In [3], Kac proved that if $p(t;\ a,\ b)$ denotes the fundamental solution of the partial differential equation:

$$\frac{\partial\varphi}{\partial t}=\frac{1}{2}\frac{\partial^2\varphi}{\partial x^2}+q(x)\varphi,$$

then

$$p(t;\ a,\ b)\,db = E_a\{e(t);\ x(t)\in db\},$$

where x is the standard Brownian motion on $\boldsymbol{R}$, and $e(t)$ is defined in (3). The associated semigroup has been called the *Kac semigroup*. It appears that Theorem 3 can be derived by Kac's original method using Laplace transforms (resolvents). This idea is due to Moerbeke, but a difficulty arises because $x\to u(x,b)$ in Theorem 3 is not necessarily bounded, so that a crucial dominated convergence required by the method appears missing. This was pointed out by Murali Rao. Although we have managed to overcome this difficulty by a moderate detour, the present approach to Theorems 3 and 4 throws more light on the "connections between probability theory and differential and integral equations" — Kac's original theme. We are indebted to Moerbeke for some stimulating discussions.

References

[1] G. Birkhoff and G. C. Rota, *Ordinary differential equations*, Ginn & Co., 1962.
[2] K. Ito and H. P. McKean, *Diffusion processes and their sample path*, Springer-Verlag, 1965.

[3] M. Kac, *On some connections between probability theory and differential and integral equations*, Proc. Second Berkeley Symposium (1951), pp. 189–215.
[4] H. P. McKean, Jr., *Stochastic integrals*, Academic Press, 1969.
[5] F. G. Tricomi, *Differential equations*, Blackie & Sons, 1961.

Added in proof: For results in R^d, $d > 1$ see K. L. Chung and K. M. Rao, *Sur la théorie du potentiel avec la fonctionnelle de Feynman-Kac,* Comp. Rend. Acad. Sci., Paris 290 (31 mars 1980).

DEPARTMENT OF MATHEMATICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

Received December 15, 1977 (1385)

Séminaire de Probabilités, XIV, 1978/1979, Lecture Notes in Math. **784**,
Springer-Verlag, Berlin (1980), 347–356.

ON STOPPED FEYNMAN-KAC FUNCTIONALS

by

Kai Lai Chung*

1. Introduction

Let $X = \{x(t),\ t \geq 0\}$ be a strong Markov process with continuous paths on $R = (-\infty, +\infty)$. Such a process is often called a diffusion. For each real b, we define the hitting time τ_b as follows:

$$(1) \qquad \tau_b = \inf\{t > 0 \mid x(t) = b\} .$$

Let P_a and E_a denote as usual the basic probability and expectation associated with paths starting from a. It is assumed that for every a and b, we have

$$(2) \qquad P_a\{\tau_b < \infty\} = 1 .$$

Now let q be a bounded Borel measurable function on R, and write for brevity

$$(3) \qquad e(t) = \exp(\int_0^t q(x(s))ds).$$

---

*Research supported in part by NSF Grant MCS77-01319, and in part by a Guggenheim fellowship in 1975-6.

This is a multiplicative functional introduced by R. Feynman and M. Kac. In this paper we study the quantity

$$u(a,b) = E_a\{e(\tau_b)\} . \tag{4}$$

Since q is bounded below, (2) implies that $u(a,b) > 0$ for every a and b, but it may be equal to $+\infty$. A fundamental property of u is given by

$$u(a,b)\, u(b,c) = u(a,c) , \tag{5}$$

valid for $a < b < c$, or $a > b > c$. This is a consequence of the strong Markov property (SMP).

2. The Results

We begin by defining two abscissas of finiteness, one for each direction.

$$\begin{aligned} \beta &= \inf\{b \in R \mid \exists\, a < b : u(a,b) = \infty\} \\ &= \sup\{b \in R \mid \forall\, a < b : u(a,b) < \infty\}; \\ \alpha &= \sup\{a \in R \mid \exists\, b > a : u(b,a) = \infty\} \\ &= \inf\{a \in R \mid \forall\, b > a : u(b,a) < \infty\}. \end{aligned} \tag{6}$$

It is possible, e.g., that $\beta = -\infty$ or $+\infty$. The first case occurs when X is the standard Brownian motion, and $q(x) \equiv 1$; for then, $u(a,b) \geqslant E_a(\tau_b) = \infty$, for any $a \neq b$.

<u>Lemma 1</u>. We have

$$\begin{aligned}\beta &= \inf\{b \in R \,|\, \forall\ a < b : u(a,b) = \infty\} \\ &= \sup\{b \in R \,|\, \exists\ a < b : u(a,b) < \infty\}\ ; \\ \alpha &= \sup\{a \in R \,|\, \forall\ b > a : u(b,a) = \infty\} \\ & \quad\ \inf\{a \in R \,|\, \exists\ b > a : u(b,a) < \infty\}\ .\end{aligned}$$

<u>Proof</u>: It is sufficient to prove the first equation above for β, because the second is trivially equivalent to it, and the equations for α follow by similar arguments. Suppose $u(a,b) = \infty$; then for $x < a < b$ we have $u(x,b) = \infty$ by (5). For $a < x < b$ we have by SMP,

$$u(x,b) \geq E_x\{e(\tau_a)\ ;\ \tau_a < \tau_b\}\ u(a,b) = \infty$$

since $P_x\{\tau_a < \tau_b\} > 0$ in consequence of (2).

The next lemma is a martingale argument. Let $\mathfrak{F}_t$ be the σ-field generated by $\{x_s,\ 0 \leq s \leq t\}$ and all null sets, so that $\mathfrak{F}_{t+} = \mathfrak{F}_t$ for $t \geq 0$; and for any optional τ let $\mathfrak{F}_\tau$ and $\mathfrak{F}_{\tau+}$ and $\mathfrak{F}_{\tau-}$ have the usual meanings.

<u>Lemma 2</u>. If $a < b < \beta$, then

$$\lim_{a \uparrow b} u(a,b) = 1\ ; \tag{7}$$

$$\lim_{b \downarrow a} u(a,b) = 1\ . \tag{8}$$

<u>Proof</u>: Let $a < b_n \uparrow b$ and consider

$$E_a\{e(\tau_b)\,|\,\mathfrak{F}(\tau_{b_n})\}\ , \qquad n \geq 1\ . \tag{9}$$

Since $b < \beta$, $u(a,b) < \infty$ and the sequence in (9) forms a martingale. As $n \uparrow \infty$, $\tau_{b_n} \uparrow \tau_b$ a.s. and $\mathfrak{F}(\tau_{b_n}) \uparrow \mathfrak{F}(\tau_n -)$. Since $e(\tau_b) \in \mathfrak{F}(\tau_b -)$, the limit of the martingale is a.s. equal to $e(\tau_b)$. On the other hand, the conditional probability in (9) is also equal to

$$E_a\{e(\tau_b)\exp(\int_{\tau_{b_n}}^{\tau_b} q(x(s))ds)\,|\,\mathfrak{F}(\tau_{b_n})\} = e(\tau_{b_n})u(b_n,b)\ .$$

As $n \uparrow \infty$, this must then converge to $e(\tau_b)$ a.s.; since $e(\tau_{b_n})$ converges to $e(\tau_b)$ a.s., we conclude that $u(b_n,b) \to 1$. This establishes (7).

Now let $\beta > b > a_n \downarrow a$, and consider

$$E_a\{e(\tau_b)\,|\,\mathfrak{F}(\tau_{a_n})\}\ , \qquad n \geq 1\ . \tag{10}$$

This is again a martingale. Although $a \to \tau_a$ is a.s. left continuous, not right continuous, for each fixed a we do have $\tau_{a_n} \downarrow \tau_a$ and $\mathfrak{F}(\tau_{a_n}) \downarrow \mathfrak{F}(\tau_a)$. Hence we obtain as before $u(a_n,b) \to u(a,b)$ and consequently

$$u(a,a_n) = \frac{u(a,b)}{u(a_n,b)} \to 1\ .$$

This establishes (8).

The next result illustrates the basic probabilistic method.

Theorem 1. The following three propositions are equivalent:

(i) $\beta = +\infty$;

(ii) $\alpha = -\infty$;

(iii) For every a and b, we have

$$u(a,b)u(b,a) \leqslant 1 . \tag{11}$$

Proof: Suppose $x(0) = b$ and let $a < b < c$. If (i) is true then $u(b,c) < \infty$ for every $c > b$. Define a sequence of successive hitting times T_n as follows (where θ denotes the usual shift operator):

$$(12)\qquad \begin{gathered} S = \begin{cases} \tau_a & \text{if } \tau_a < \tau_c , \\ \infty & \text{if } \tau_c < \tau_a ; \end{cases} \\ T_0 = 0 , \qquad T_1 = S , \\ T_{2n} = T_{2n-1} + \tau_b \circ \theta_{T_{2n-1}} , \qquad T_{2n+1} = T_{2n} + S \circ \theta_{T_{2n}} , \end{gathered}$$

for $n \geqslant 1$. Define also

$$N = \min\{n \geqslant 0 \,|\, T_{2n+1} = \infty\} . \tag{13}$$

It follows from $P_b\{\tau_c < \infty\} = 1$ that $0 \leqslant N < \infty$ a.s. For $n \geqslant 0$, we have

352

$$E_b\{e(\tau_c)\ ;\ N = n\} = E_b\{\exp(\sum_{k=0}^{2n} \int_{T_k}^{T_{k+1}} q(x(s))ds)\}$$

(14)

$$= E_b\{e(\tau_a)\ ;\ \tau_a < \tau_c\}^n\ E_a\{e(\tau_b)\}^n\ E_b\{e(\tau_c)\ ;\ \tau_c < \tau_a\}\ .$$

Since the sum of the first term in (14) over $n \geq 0$ is equal to $u(b,c) < \infty$, the sum of the last term in (14) must converge. Thus we have

$$E_b\{e(\tau_a)\ ;\ \tau_a < \tau_c\}\ u(a,b) < 1\ . \tag{15}$$

Letting $c \to \infty$ we obtain (11). Hence $u(b,a) < \infty$ for every $a < b$ and so (ii) is true. Exactly the same argument shows that (ii) implies (iii) and so also (i).

We are indebted to R. Durrett for ridding the next lemma of a superfluous condition.

Lemma 3. Given any $a \in R$ and $Q > 0$, there exists an $\varepsilon = \varepsilon(a,Q)$ such that

$$E_a\{e^{Q\sigma_\varepsilon}\} < \infty \tag{16}$$

where

$$\sigma_\varepsilon = \inf\{t > 0\,|\,x(t) \notin (a - \varepsilon,\ a + \varepsilon)\}\ .$$

Proof: Since X is strong Markov and has continuous paths, there is no "stable" point. This implies $P_a\{\sigma_\varepsilon \geq 1\} \to 0$ as $\varepsilon \to 0$ and so there exists ε such that

(17) $$P_a\{\sigma_\varepsilon \geq 1\} < e^{-(Q+1)} .$$

Now σ_ε is a terminal time, so $x \to P_x\{\sigma_\varepsilon \geq 1\}$ is an excessive function for the process X killed at σ_ε. Hence by standard theory it is finely continuous. For a diffusion under hypothesis (2) it is clear that fine topology coincides with the Euclidean. Thus $x \to P_x\{\sigma_\varepsilon \geq 1\}$ is in fact continuous. It now follows that we have, further decreasing ε if necessary:

(18) $$\sup_{|x-a|<\varepsilon} P_x\{\sigma_\varepsilon \geq 1\} < e^{-(Q+1)} .$$

A familiar inductive argument then yields for all $n \geq 1$.

(19) $$P_a\{\sigma_\varepsilon \geq n\} < e^{-n(Q+1)}$$

and (16) follows.

Lemma 4. For any $a < \beta$ we have

(20) $$u(a,\beta) = \infty ;$$

for any $b > \alpha$ we have $u(b,\alpha) = \infty$.

Proof: We will prove that if $u(a,b) < \infty$, then there exists $c > b$ such that $u(b,c) < \infty$. This implies (20) by Lemma 1, and the second assertion is proved similarly.

Let $Q = \|q\|$. Given b we choose a and b so that $a < b < d$ and

$$(21) \qquad E_b\{e^{Q(\tau_a \wedge \tau_d)}\} < \infty .$$

This is possible by Lemma 3. Now let $b < c < d$; then as $c \downarrow b$ we have

$$(22) \quad E_b\{e(\tau_a);\ \tau_a < \tau_c\} \leqslant E_b\{e^{Q(\tau_a \wedge \tau_d)};\ \tau_a < \tau_c\} \to 0$$

because $P_b\{\tau_a < \tau_c\} \to 0$. Hence there exists c such that

$$(23) \qquad E_b\{e(\tau_a);\ \tau_a < \tau_c\} < \frac{1}{u(a,b)} .$$

This is just (15) above, and so reversing the argument there, we conclude that the sum of the first term in (14) over $n \geqslant 0$ must converge. Thus $u(b,c) < \infty$, as was to be shown.

To sum up:

<u>Theorem 2</u>. The function $(a,b) \to u(a,b)$ is continuous in the region $a \leqslant b < \beta$ and in the region $\alpha < b \leqslant a$. Furthermore, extended continuity holds in $a \leqslant b \leqslant \beta$ and $\alpha \leqslant b \leqslant a$, except at (β,β) when $\beta < \infty$, and at (α,α) when $\alpha > -\infty$.

<u>Proof</u>: To see that there is continuity in the extended sense at (a,β), where $a < \beta$, let $a < b_n \uparrow \beta$. Then we have by Fatou's lemma

$$\underline{\lim}_{n\to\infty} u(a,b_n) \geq E_a\{\lim_{n\to\infty} e(\tau_{b_n})\} = E_a\{e(\tau_\beta)\} = u(a,\beta) = \infty .$$

If $\beta < \infty$, then $u(\beta,\beta) = 1$ by dèfinition, but $u(a,\beta) = \infty$ for all $a < \beta$; hence u is not continuous at (β,β). The case for α is similar.

3. <u>The Connections</u>

Now let X be the standard Brownian motion on R and q be bounded and continuous on R.

<u>Theorem 3</u>. Suppose that $u(x,b) < \infty$ for some, hence all, $x < b$. Then $u(\cdot,b)$ is a solution of the Schrödinger equation:

$$\frac{1}{2}\varphi'' + q\varphi = 0$$

in $(-\infty,b)$ satisfying the boundary condition

$$\lim_{x\to b} \varphi(x) = 1.$$

There are several proofs of this result. The simplest and latest proof was found a few days ago while I was teaching a course on Brownian motion. This uses nothing but the theorem by H. A. Schwarz on generalized second derivative and the continuity of $u(\cdot,b)$ proved in Theorem 2. It will be included in a projected set of lecture notes. An older proof due to Varadhan and using Ito's calculus and martingales will be published

elsewhere. An even older unpublished proof used Kac's method of Laplace transforms of which an incorrect version (lack of domination!) had been communicated to me by an *ancien collègue*.

But none of these proofs will be given here partly because they constitute excellent exercises for the reader, and partly because the results have recently been established in any dimension (for a bounded open domain in lieu of $(-\infty, b)$), in collaboration with K. M. Rao. These are in the process of consolidation and extension.

I am indebted to Pierre van Moerbeke for suggesting the investigation in this note. The situation described in Theorem for the case of Brownian motion apparently means the absence of "bound states" in physics!

Ann. Inst. Fourier, Grenoble
30, 3 (1980), 167-198

A NEW SETTING FOR POTENTIAL THEORY (Part 1)

by K.L. CHUNG and K. MURALI RAO (*)

Introduction.

In Hunt's theory of Markov processes certain duality assumptions are made to generalize well known classical potential theoretic results such as F. Riesz representation theorem, uniqueness, existence of equilibrium potential etc. A standard treatment developed by several subsequent authors can be found in [1]. In a different direction, it was shown in [2] under simple analytic conditions that the equilibrium measure is inherently linked to the last exit distribution of the process. It thus appears feasible that this last result, namely on "equilibrium principle" may be made the starting point to which other major results are related.

In this paper we exploit the line of thought in [2] to derive some of these major results under the same set of conditions as in [2]. It turns out that these sets of conditions automatically imply the existence of a dual. However, this will not be proved here.

In § 1 we collect a number of simple consequences of our basic assumptions. In § 2 we construct a version of the potential density with "point supports" namely which is "harmonic off its pole". This is a general result not dependent on the specific conditions of the present paper. This good version of the potential density turns out

---

(*) Research of both authors was supported in part by NSF grant MCS77-01319. Some of the work was done by the first named author when he was a Guggenheim fellow during 1976-77.

to be indispensable in settling the question of uniqueness of representing measure in § 3 and § 4. § 4 also contains the Riesz representation theorem. These results automatically lead to a dual semigroup which is used to prove Hunt's Hypothesis (B) under mild supplementary conditions in § 5. Further development of these results is relegated to Part II.

1.

The basic assumptions are the same as in [2], viz. :

(i) The underlying process X is a Hunt process on a locally compact Hausdorff space E with countable base, which is transient in the following sense: for each compact K and every x we have

$$\lim_{t\to\infty} P^x\{T_K \circ \theta_t\} = 0. \tag{1}$$

(ii) The potential kernel is $U(x, dy) = u(x, y)\,\xi(dy)$ where ξ is a Radon measure and the potential density function u has the following properties:

(iia) $\forall x : y \longrightarrow \dfrac{1}{u(x, y)}$ is finite continuous

(iib) $u(x, y) = +\infty$ if and only if $x = y$.

To save constant repetition we shall fix the usage of certain symbols and terms below (unless explicitly contravened), as follows:

A is a Borel set, written also as $A \in \mathfrak{B}$;
D is an open set with compact closure, not empty;
G is an open set, not empty;
K is a compact set, not empty;

a function such as f or f_n is positive Borel measurable; the support of a function f or a measure μ is denoted by $\lfloor\!\underline{f}$ or $\lfloor\!\underline{\mu}$;

$P_t u(x, y) = \int P_t(x, dz)\, u(z, y)$, $P_A u(x, y) = \int P_A(x, dz)\, u(z, y)$,

where (P_t) is the (Borelian) semigroup of the process

(X_t) ; $P_A f(x) = E^x\{f(X_{T_A})\,;\, T_A < \infty\}$, $T_A = \inf\{t > 0 : X_t \in A\}$;

s is superaveraging iff $s \geqslant P_t s$ for every t; the excessive regularization of s is denoted by $\underline{s} = \lim_{t\to 0} P_t s$; s is excessive at x iff $s(x) = \underline{s}(x)$;

a potential s is an excessive function such that $\lim\limits_{K\uparrow E} P_{K^c}s = 0$, ξ a.e.;

$A_n \downarrow\downarrow A$ means $A_n \supset \overline{A}_{n+1}$ for all n and $\bigcap_n A_n = A$;

$\mu_1 < \mu_2$ means μ_1 is absolutely continuous with respect to μ_2. A measure is diffuse iff it does not charge any singleton.

We list here some simple consequences of the basic assumptions.

PROPOSITION 1. – *There exists $h > 0$ everywhere such that $Uh \leqslant 1$ everywhere.*

Proof. – This may be known, but observe that we make no assumption on lower semi-continuity of $x \longrightarrow u(x, y)$. Here is a proof. Let $D_k \uparrow E$, there exists t_k such that

$$P_{D_k}1(x) - P_{t_k}P_{D_k}1(x) > 0 \text{ on } A_k \subset D_k \text{ with } \xi(D_k - A_k) < \frac{1}{2^k}.$$

The series

$$\sum_{k=1}^{\infty} \frac{1}{2^k t_k} (P_{D_k}1 - P_{t_k}P_{D_k}1)$$

converges everywhere and is strictly positive ξ-a.e. Put h to equal the sum where it is > 0; $= 1$ elsewhere.

It is convenient to put

$$\xi_0(dy) = h(y)\,\xi(dy). \tag{2}$$

For each x, the measure $u(x, y)\,\xi_0(dy)$ is then finite, whereas $u(x, y)\,\xi(dy)$ may not be a Radon measure. Since ξ_0 is equivalent to ξ, this is convenient for applications of Fubini's theorem.

PROPOSITION 2. – *ξ is a diffuse measure.*

Proof. – This follows from $\infty > U1_{\{x\}}(x) \geqslant u(x, x)\,\xi(\{x\})$.

PROPOSITION 3. – *$\forall y : u(\cdot, y)$ is superaveraging.*

Proof. – Since $P_t Uf \leqslant Uf$ for any f it follows that $\forall x$, $\exists N_x$ with $\xi(N_x) = 0$ such that $P_t u(x, y) \leqslant u(x, y)$ for $y \notin N_x$. Let $y_n \notin N_x$, $y_n \longrightarrow y$, then $\forall z : u(z, y_n) \longrightarrow u(z, y)$. Hence by Fatou's lemma,

$$P_t u(x, y) \leqslant \varliminf_n P_t u(x, y_n) \leqslant \varliminf_n u(x, y_n) = u(x, y). \qquad \square$$

For each y, we write $\underline{u}(\cdot\,,y)$ for the excessive regularization of u ; $\underline{U}f = \int u(\cdot\,,y)f(y)\,dy$, $\underline{U}\mu = \int \underline{u}(\cdot\,,y)\,\mu(dy)$.

PROPOSITION 4. – *We have*

$$\forall f : \underline{U}f = Uf. \tag{3}$$

If $s = U\mu$ *where* μ *is any measure, then*

$$\underline{s} = \underline{U}\mu . \tag{4}$$

Proof. – (3) is true because both members are excessive and they are equal ξ-a.e. by Fubini; (4) contains (3) and is proved by Fubini and Fatou.

PROPOSITION 5. – *If* s *is excessive, then* $\exists f_n$ *such that* $f_n \leqslant n^2$, $Uf_n \leqslant n$, *and* $Uf_n \uparrow s$.

Proof. – This is well known but we indicate the proof. Let $K_n \uparrow\uparrow E$, $s_n = s \wedge (n P_{K_n} 1)$ and $f_n = n[s_n - P_{1/n}\, s_n]$.

The next two propositions are proved in [2], reviewed here for orientation and some quick applications.

PROPOSITION 6. – *For each* K , $\exists$ *a Radon measure* μ *such that*

$$P_K 1 = U\mu . \tag{5}$$

It is important to observe that the proof in [2] does not establish that $\lfloor\!\underline{\mu} \subset K$. In fact it shows that

$$\int_A u(x\,,y)\,\mu(dy) = P^x\{\gamma_K > 0\,;\, X(\gamma_K -) \in A\} \tag{6}$$

where $\gamma_K = \sup\{t > 0 : X_t \in K\}$. If there is a jump at γ_K , it is possible that $X(\gamma_K -) \notin K$. However, (6) does establish the next proposition.

PROPOSITION 7. – *For each* D , $\exists\mu$ *with* $\lfloor\!\underline{\mu} \subset \overline{D}$ *such that*

$$P_D 1 = U\mu . \tag{7}$$

One of our principal results below is to prove that there exists μ with $\lfloor\!\underline{\mu} \subset K$ and $\mu(Z) = 0$ (see (12) of § 3), for which (5) is true. This turns out to be equivalent to Hunt's Hypothesis (B) and

will be proved under an additional assumption. For the moment we note some consequences of the two preceding propositions.

PROPOSITION 8. – *If* s *is excessive and* $s \not\equiv \infty$, *then the set* $\{s = \infty\}$ *is polar, in particular* $s < \infty$ ξ-a.e.

Proof. – Let $K \subset \{s = \infty\}$, and $s(x_0) < \infty$. Then it is well known that $P_K 1(x_0) = 0$, namely by (5): $\int u(x_0, y)\, \mu(dy) = 0$. Since $u(x, y) > 0$ for all (x, y) by basic assumption (iia) this implies $\mu \equiv 0$ and so the preceding equation holds for any x instead of x_0. Thus K is polar, and so is $\{s = \infty\}$.

COROLLARY. – *If* $U\mu \not\equiv \infty$, *then* μ *is a Radon measure; moreover,* $U\mu < \infty$ ξ-a.e.

Proof. – Let $s = U\mu$. If $s(x_0) < \infty$, $\forall K$:

$$\infty > \int_K u(x_0, y)\, \mu(dy).$$

Since $\inf_{y \in K} u(x_0, y) > 0$ by (iia), it follows that $\mu(K) < \infty$. Hence μ is Radon. Now write $s = U\mu$, then $\underline{s} \not\equiv \infty$, hence by the proposition $\underline{s} < \infty$ ξ-a.e. Since $s = \underline{s}$ ξ-a.e., we have $s < \infty$ ξ-a.e.

PROPOSITION 9. – *If* s *is excessive and* $s \not\equiv 0$, *then* $s > 0$ *everywhere.*

Proof. – By Proposition 5, $\exists f_n$ such that $Uf_n \uparrow s$. Hence if $s(x_0) > 0$, then $\int u(x_0, y) f_n(y)\, \xi(dy) > 0$ for large n. Since $u > 0$, the same is then true if x_0 is replaced by any x. Hence $s(x) > 0$.

PROPOSITION 10. – *Each singleton is a polar set.*

Proof. – Let $D_n \downarrow\downarrow \{x_0\}$; by Proposition 7, $P_{D_n} 1 = U\mu_n$ where $|\mu_n \subset \overline{D}_n$. For each K and any fixed x:

$$1 \geqslant \int_K u(x, y)\, \mu_n(dy) \geqslant \inf_{y \in K} u(x, y)\, \mu_n(K).$$

Hence $\{\mu_n\}$ is vaguely bounded and a subsequence converges vaguely to μ, which must be supported by $\{x_0\}$, namely $\mu = \lambda \delta_{x_0}$ for

some $\lambda \geqslant 0$. Since for each x, $u(x,\cdot)$ is extended continuous, we have $\lambda u(x, x_0) = U\mu(x) \leqslant \underline{\lim}_n U\mu_n(x) \leqslant 1$. For $x = x_0$ this yields $\lambda = 0$ by basic assumption (iib); thus $\mu \equiv 0$. Now let $x_1 \notin \overline{D}_1$, then $y \longrightarrow u(x_1, y)$ is finite continuous in $\overline{D}_1$. Hence we obtain

$$\lim_n P_{D_n} 1(x_1) = \lim_n \int_{\overline{D}_1} u(x_1, y)\, \mu_n(dy) = \int_{\overline{D}_1} u(x_1, y)\, \mu(dy) = 0$$

by vague convergence, and consequently $P_{\{x_0\}} 1(x_1) = 0$. By Proposition 9, $P_{\{x_0\}} 1 \equiv 0$, namely $\{x_0\}$ is polar.

The following lemma is known (see [1 ; p. 84]) but we supply a proof for the sake of completeness.

LEMMA. – *Let* $K_n \uparrow E$ *and* $T_n = T_{K_n^c}$, or $T_n = n$. *Then for any excessive function* f *we have*

$$P_{T_n} f \downarrow g \tag{8}$$

where $g = \underline{g}$ *on* $\{g < \infty\}$.

Proof. – The case $T_n = n$ is easy; so we treat only the case $T_n = T_{K_n^c}$. It is clear from (8) that g is superaveraging. Fix an x such that $g(x) < \infty$, then for $n \geqslant n_0(x)$ we have

$$\begin{aligned} \infty > P_{T_n}(f(x) \geqslant P_t P_{T_n} f(x) \geqslant E^x\{f(X(t + T_n \circ \theta_t))\, ;\, t < T_n\} \\ = E^x\{f(X(T_n))\, ;\, t < T_n\}. \end{aligned}$$

It follows by subtraction that

$$P_{T_n} f(x) - P_t P_{T_n} f(x) \leqslant E^x\{f(X(T_n))\, ;\, T_n \leqslant t\}. \tag{9}$$

For fixed t, the second member of (9) is decreasing in n by the supermartingale inequality. Hence for $n \geqslant k$, (9) is true when on the right side n is replaced by k. Now let $n \longrightarrow \infty$ on the left side and use domination to obtain for every k

$$g(x) - P_t g(x) \leqslant E^x\{f(X(T_k))\, ;\, T_k \leqslant t\}. \tag{10}$$

For the fixed x there exists k such that $P^x\{T_k > 0\} = 1$. For this k in (10), as $t \downarrow 0$ the right member of (10) decreases to zero by domination; $P_{T_k} f(x) < \infty$. Hence $g(x) = \underline{g}(x)$. □

PROPOSITION 11. – *If s is potential, then*

$$\lim_{K \uparrow E} P_{K^c} s(x) = 0 \tag{11}$$

if $s(x) < \infty$. Hence the set of x for which (11) does not hold is polar.

Proof. – Apply the Lemma above to $\{P_{K_n^c} s\}$. Then we have by definition $g = \lim_n P_{K_n^c} s = 0$, ξ-a.e. Hence $\underline{g} \equiv 0$ and so by the Lemma, $g = 0$ on $\{g < \infty\} \supset \{s < \infty\}$. Standard arguments show that if $g(x) = 0$, then (11) is true.

This proposition will be crucial in the proof of Theorem 2, (δ) below.

2.

In this section we give a general construction of a good version of a given potential density function u. Of the latter we assume that

(a) $\forall y : u(\cdot, y)$ is excessive and finite ξ-a.e.

For the underlying process it is sufficient that it be a transient standard process satisfying the condition:

(b) every singleton is a polar set.

A function v measurable $\mathfrak{B} \times \mathfrak{B}$ is a version of u iff for every x and every f we have $\int v(x, y) f(y) \xi(dy) = \int u(x, y) f(y) \xi(dy)$.

THEOREM 1. – *Under the conditions stated above, there is a version w of u which has the following properties: for each y, $w(\cdot, y)$ is excessive; and for each (open) G we have*

$$P_G w(x, y) = w(x, y) \quad \textit{for all} \quad y \in G. \tag{1}$$

We shall refer to (1) as the "round" property of w.

Proof. – Let $\{B_n\}$ be a countable base of open sets of E such that each member of the sequence is repeated infinitely many times. Put $u_0(x, y) = u(x, y)$ for all (x, y); and define inductively for $n \geqslant 0$:

$$u_{n+1}(x, y) = \begin{cases} P_{B_{n+1}} u_n(x, y) & \text{if} \quad y \in B_{n+1}, \\ u_n(x, y) & \text{if} \quad y \notin B_{n+1}. \end{cases}$$

For each y, $u_n(\cdot, y)$ is excessive by assumption (a), and

$$u_n(x, y) \geqslant u_{n+1}(x, y)$$

for all (x, y). Let $u_\infty(x, y) = \lim_n \downarrow u_n(x, y)$. Then $u_\infty(\cdot, y)$ is superaveraging.

Recall the measure ξ_0 in (2) of § 1. We have $P_B U(1_B h) = U(1_B h)$ for open B, namely

$$\int_B P_B u(x, y)\, \xi_0(dy) = \int_B u(x, y)\, \xi_0(dy) < \infty. \tag{2}$$

(ξ_0 is used in lieu of ξ to ensure finiteness above.) Hence by Fubini there exists N with $\xi_0(N) = 0$ such that

$$P_B u(x, y) = u(x, y), \qquad \text{for all} \quad y \notin N \quad \text{and} \quad y \in B \tag{3}$$

first for ξ_0-a.e. x, then for all x because both members of (3) are excessive, and ξ_0 is a reference measure as well as ξ. Applying (3) with $B = B_1$, we obtain $u_1(x, y)\, \xi_0(dy) = u(x, y)\, \xi_0(dy)$ and hence by induction $u_n(x, y)\, \xi_0(dy) = u(x, y)\, \xi_0(dy)$. Since both members above are Radon measures, it follows by monotone convergence that

$$u_\infty(x, y)\, \xi_0(dy) = u(x, y)\, \xi_0(dy). \tag{4}$$

Thus u_∞ is a version of u, but it may not be excessive.

For each y let us put $F_y = \{x \mid u_\infty(x, y) < \infty\} \backslash \{y\}$. Then $\xi(F_y^c) = 0$ by assumptions (a) and (b). Let B be one of the sequence $\{B_n\}$ and $y \in B$. Then by construction there exists a sequence $n_k \longrightarrow \infty$ such that $B_{n_k} = B$ and so $P_B u_{n_k - 1}(x, y) = u_{n_k}(x, y)$. It follows by monotone convergence with initial finiteness that

$$P_B u_\infty(x, y) = u_\infty(x, y) \quad \text{for all} \quad x \in F_y \quad \text{and} \quad y \in B. \tag{5}$$

Next, if $x \in F_y$, then $u_n(x, y) < \infty$ for all large n. Since $u_n(\cdot, y)$ is excessive, $u_n(X_t, y)$ is a right continuous supermartingale under P^x, for large n. Hence by Doob's stopping theorem

$$E^x\{u_n(X_t, y)\,;\; t \leqslant T_B\} \geqslant E^x\{u_n(X_{T_B}, y)\,;\; t \leqslant T_B\}. \tag{6}$$

Since for large n $E^x\{u_n(X_t, y)\} \leqslant u_n(x, y) < \infty$, we can let $n \longrightarrow \infty$ in (6) to obtain

$$\begin{aligned} E^x\{u_\infty(X_t, y) ; t \leqslant T_B\} &\geqslant E^x\{u_\infty(X_{T_B}, y) ; t \leqslant T_B\} \qquad (7)\\ &= P_B u_\infty(x, y) - E^x\{u_\infty(X_{T_B}, y) ; t > T_B\}. \end{aligned}$$

The last expectation in (7) is bounded by $E^x\{u_n(X_{T_B}, y) ; t > T_B\}$. Since $x \neq y$, we may choose B so that $x \notin \overline{B}$ so that $P^x\{T_B > 0\} = 1$. The expectation above then converges to zero as $t \longrightarrow 0$ because

$$E^x\{u_n(X_{T_B}, y) ; T_B < \infty\} = P_B u_n(x, y) \leqslant u_n(x, y) < \infty$$

for large n. Going back to (7), we see that if $x \in F_y$, then

$$\begin{aligned} \varliminf_{t \downarrow 0} P_t u_\infty(x, y) &\geqslant \varliminf_{t \downarrow 0} E^x\{u_\infty(X_t, y) ; t \leqslant T_B\}\\ &\geqslant P_B u_\infty(x, y) = u_\infty(x, y). \end{aligned}$$

Thus $u_\infty(\cdot, y)$ is excessive at such an x, namely

$$\underline{u}_\infty(x, y) = u_\infty(x, y), \quad \text{for all} \quad x \in F_y. \qquad (8)$$

In addition to equation (5), we have if B is any member of $\{B_n\}$:

$$P_B u_\infty(x, y) \leqslant u_\infty(x, y) \quad \text{for all} \quad (x, y), \qquad (9)$$

because this is true when u_∞ is replaced by u_n, which implies (9) itself by Fatou's lemma. Now if $x \in F_y$, then the quantities in (9) are finite, and consequently in conjunction with $P_B(x, \{y\}) = 0$ due to the polarity of $\{y\}$, we have $P_B(x, F_y^c) = 0$. This relation and (8) imply that $P_B u_\infty(x, y) = P_B \underline{u}_\infty(x, y)$. Since $\underline{u}_\infty(\cdot, y)$ is excessive, we now have

$$u_\infty(x, y) \geqslant \underline{u}_\infty(x, y) \geqslant P_B \underline{u}_\infty(x, y) = P_B u_\infty(x, y) = u_\infty(x, y)$$

for all $y \in B$ and all $x \in F_y$. Since $\xi(F_y^c) = 0$, we conclude that

$$P_B \underline{u}_\infty(x, y) = \underline{u}_\infty(x, y) \quad \text{for all} \quad y \in B \qquad (10)$$

first for ξ-a.e. x, then for all x. The validity of (10) for all members of a base of the topology implies its validity for every open set B.

Finally, the proof of Proposition 4 of § 1 shows that $\underline{u}_\infty$ as well as u_∞ is a version of u. This is the w claimed in Theorem 1. □

To apply Theorem 1 to the case under consideration in this paper, we must start with $\underline{u}$ instead of u because of condition (a). Note that by (3) of § 1, $\underline{u}$ is a potential density of the given process. Proposition 10 of § 1 supplies the condition (b) required for Theorem 1.

We shall refer to the w just constructed from $\underline{u}$ as *the round version* of u. It will play a key role in what follows.

3.

In this section we give the principal convergence theorem for potentials of measures. It is an extension of the main result in [9] but will soon be strengthened to include a new feature relative to the round version w constructed in § 2.

THEOREM 2. – *Let* $\{\mu_n\}$ *be a sequence of diffuse measures, satisfying conditions (a), (b) and either* (c_1) *or* (c_2) *below:*

(a) $\forall n : U\mu_n \leqslant \sigma$ *where* σ *is excessive and* $\not\equiv \infty$;

(b) $\lim_n U\mu_n = s$ *everywhere;*

(c_1) $\forall n : \underline{|\mu_n}$ *is contained in a fixed compact;*

(c_2) $\forall n : \mu_n \prec \xi$ *and* σ *is a potential.*

Then there exists a subsequence $\{n_j\}$ *and a Radon measure* μ *such that*

(α) μ_{n_j} *converges vaguely to* μ;

(β) *at each* x *where* $\sigma(x) < \infty$ *and* $s(x) = \underline{s}(x) < \infty$, *the measures* $u(x, y)\mu_{n_j}(dy)$ *converge weakly to* $u(x, y)\mu(dy)$, *in particular* $s(x) = U\mu(x)$;

(γ) *if the* s *in condition (b) is an excessive function* $\not\equiv \infty$, *then* $s = U\mu$ *everywhere.*

Proof. – The proof is essentially the same as that in [9], but the basic steps will be sketched.

Let $\sigma(x_0) < \infty$, then by (a), for each K,

$$\sigma(x_0) \geqslant [\inf_{y \in K} u(x_0, y)]\, \mu_n(K)$$

where the infimum is strictly positive. This implies (α). We shall write μ_n for μ_{n_j} below for simplicity. Put

$$L_n(x, dy) = u(x, y)\,\mu_n(dy)\,; \quad L_n(x, f) = \int L_n(x, dy) f(y)\,.$$

For each x for which $\sigma(x) < \infty$, we have by (a) $\infty > \sigma(x) \geqslant L_n(x, 1)$. Hence there exists $\{n_j\}$ and $L(x, \cdot)$ such that

$$L_{n_j}(x, \cdot) \longrightarrow L(x, \cdot) \quad \text{vaguely}. \tag{1}$$

Let φ be continuous with compact support, then since $\forall n$: $\mu_n(\{x\}) = 0$ by hypothesis and $u(x, y) < \infty$ for $y \neq x$, we have

$$\begin{aligned} \int \varphi(y)\, \mu(dy) &= \lim_j \int \varphi(y)\, \mu_{n_j}(dy) \qquad (2) \\ &= \lim_j \int \frac{\varphi(y)}{u(x, y)}\, L_{n_j}(x, dy) = \int \frac{\varphi(y)}{u(x, y)}\, L(x, dy). \end{aligned}$$

Thus we have for each x such that $\sigma(x) < \infty$:

$$\mu(dy) = \frac{L(x, dy)}{u(x, y)}\,. \tag{3}$$

This is true of any vaguely convergent subsequence of $L_n(x, \cdot)$, hence by (3) any two vague limits coincide off $\{x\}$. Now under condition (c_1) the vague convergence is also weak convergence, namely with $L_n(x, 1) \longrightarrow L(x, 1)$. On the other hand, under condition (c_2) we may write $\mu_n(dy) = f_n(y)\, \xi(dy)$, and so for each K

$$\int_{K^c} u(x, y)\, \mu_n(dy) \leqslant U(f_n 1_{K^c})(x) = P_{K^c} U(f_n 1_{K^c})(x) \leqslant P_{K^c}\, \sigma(x). \tag{4}$$

By Proposition 11 of § 1, if $\sigma(x) < \infty$, then the last term in (4) decreases to zero as $K \uparrow E$. Consequently we have

$$L_n(x, \cdot) \longrightarrow L(x, \cdot) \quad \text{weakly} \tag{5}$$

and $L(x, 1) = s(x)$ by condition (b).

Let $F = \{y : \sigma(y) < \infty\}$. Then F^c is a polar set and we have proved under (c_1) or (c_2) that (5) is true for all $x \in F$. It follows from (3) that $\mu(\{x\}) = 0$ and

$$L(y, \{x\}) = 0 \quad \text{if} \quad x \in F,\ y \in F,\ x \neq y. \tag{6}$$

The limit s in (b) is superaveraging by Fatou. We are going to show that if $s(x) = \underline{s}(x) < \infty$, then

$$L(x, \{x\}) = 0\,. \tag{7}$$

Let g be continuous and $0 \leqslant g \leqslant 1$. Using (5) we see that $L(\cdot, 1 - g)$ is superaveraging, hence

$$L(x, g) - P_t L(x, g) \leqslant L(x, 1) - P_t L(x, 1).$$

The right member converges to zero as $t \longrightarrow 0$ by hypothesis. Taking a sequence $g_n \downarrow 1_{\{x\}}$, we obtain

$$L(x, \{x\}) \leqslant \epsilon + P_t L(x, \{x\}) \tag{8}$$

for any $\epsilon > 0$ and sufficiently small $t > 0$. Now $F^c \cup \{x\}$ is a polar set by Proposition 10 of § 1, hence $P_t(x, \cdot)$ does not charge it and so the last term in (8) is equal to zero by (6). Thus (7) follows from (8). We can now conclude from (3) that

$$L(x, dy) = u(x, y)\,\mu(dy). \tag{9}$$

This and (5) establish the conclusion (β). Integrating, we obtain

$$s(x) = U\mu(x) \tag{10}$$

except possibly for a polar set. Under the hypothesis in (γ), this implies $s = \underline{U}\mu$ everywhere. But the lower semicontinuity of $u(x, \cdot)$ for each x and the vague convergence of μ_n to μ also implies that $s \geqslant U\mu$ everywhere. Thus $s = U\mu$ as asserted. □

COROLLARY 1. – *If s is excessive and $\not\equiv \infty$, then $P_D s = U\mu$, where $\lfloor\mu \subset \overline{D}$.*

Proof. – By a theorem of Hunt's [3], there exists f_n with $\lfloor f_n \subset D$ such that $Uf_n \uparrow P_D s$. Hence (a), (b) and (c_1) are satisfied with s and σ both equal to $P_D s$ here, since $P_D s < \infty$ ξ-a.e. by Proposition 8 of § 1. It is trivial that $\lfloor\mu \subset \overline{D}$ by vague convergence.

COROLLARY 2. – *If s is a potential, then there is a Radon measure μ such that $s = U\mu$.*

Proof. – By Proposition 5 of § 1, (a), (b) and (c_2) are satisfied if $\sigma = s$. Note that a potential is necessarily $\not\equiv \infty$, hence $< \infty$ ξ-a.e.

COROLLARY 3. – *For each K, we have $P_K 1 = \underline{U}\mu$, where $\lfloor\mu \subset K$; also, $P_K 1 = U\mu$ on K^c.*

Proof. – Let $D_n \downarrow\downarrow K$; by Corollary 2, we have $P_{D_n} 1 = U\mu_n$, where $\lfloor\mu_n \subset \overline{D}_n$. Put $s = \lim_n \downarrow P_{D_n} 1$ and apply Theorem 2 under (c_1) to obtain a subsequence $\{\mu_{n_j}\}$ converging vaguely to μ, such that $s = U\mu$ on $\{s = \underline{s}\}$. Then $\lfloor\mu \subset \bigcap_n \overline{D}_n = K$ and $P_K 1 = \underline{s} = \underline{U}\mu$.

If $x \notin K$, then $u(x,\cdot)$ is continuous in $\overline{D}_n$ for large n; on the other hand, $s(x) = \underline{s}(x)$. This establishes the second assertion of the lemma.

We can now prove a key property of the round version w of u.

THEOREM 3. – *For any* y : *either* $w(\cdot, y) \equiv u(\cdot, y)$ *or* $w(\cdot, y) \equiv 0$.

Proof. – Fix y and let $D_n \downarrow\downarrow \{y\}$. Since $w(\cdot, y)$ is excessive and $< \infty$ ξ-a.e., we have by Corollary 1 to Theorem 2

$$P_{D_n} w(\cdot, y) = U\mu_n, \quad \text{where } \underline{\mu_n} \subset \overline{D}_n. \tag{11}$$

But the left member of (11) is just $w(\cdot, y)$. Hence Theorem 2 is applicable to the sequence $\{U\mu_n\}$ under condition (c_1), and we conclude that there is a subsequence μ_{n_j} converging vaguely to some μ such that $w(\cdot, y) = U\mu$. But μ must have support in $\bigcap_n \overline{D}_n = \{y\}$, thus $\exists\lambda : 0 \leqslant \lambda < \infty$, such that $w(\cdot, y) = u(\cdot, y)\lambda$. If $\lambda = 0$, then $w(\cdot, y) \equiv 0$. If $\lambda > 0$, then $u(\cdot, y) = \dfrac{w(\cdot, y)}{\lambda}$, and consequently $u(\cdot, y)$ is excessive and furthermore for any $G \ni y$, $P_G u(\cdot, y) = u(\cdot, y)$. The construction of Theorem 1 then yields successively $u_n(\cdot, y) \equiv u(\cdot, y)$, $u_\infty(\cdot, y) \equiv u(\cdot, y)$ and finally $w(\cdot, y) \equiv u(\cdot, y)$. □

We now introduce the all important exceptional set below:

$$Z = \{y : w(\cdot, y) \equiv 0\} = \{y : w(\cdot, y) \not\equiv u(\cdot, y)\}. \tag{12}$$

Clearly $Z \in \mathfrak{B}$. Note that $U\mu = W\mu$ if and only if $\mu(Z) = 0$.

PROPOSITION 12. – *We have* $\xi(Z) = 0$.

Proof. – Since w is a potential density, we have

$$0 = \int_Z w(x, y)\, \xi_0(dy) = \int_Z u(x, y)\, \xi_0(dy)$$

where ξ_0 is defined in (2) of § 1. But $u > 0$ everywhere, hence $\xi_0(Z) = 0$ which is the same as $\xi(Z) = 0$. □

Theorem 2 under condition (c_2) was stated in a restricted way because we needed its Corollary 1 to prove Theorem 3. We can now state Theorem 2 in a more complete form as follows.

THEOREM 2 (continued). – *If we impose the additional condition that* $\mu_n(Z) = 0$ *for all* n, *then in condition* (c_2) *we may remove the assumption that* $\mu_n \prec \xi$, *and the conclusions* (α), (β), (γ) *still hold; moreover, we have*

(δ) *for the* μ *in* (γ) *we have* $\mu(Z) = 0$.

Proof. – The proof of (α) requires no change. In the proof of (β), under condition (c_2), the inequality (4) is replaced as follows. Let $K_n \uparrow E$; since $\mu_n(Z) = 0$, we have

$$\int_{K_n^c} u(x, y)\, \mu_n(dy) = \int_{K_n^c} w(x, y)\, \mu_n(dy). \tag{13}$$

By the round property of w, $P_{K_n^c} w(\cdot, y) = w(\cdot, y)$ for all $y \in K_n^c$. Hence the second member of (13) does not exceed

$$P_{K_n^c}\left[\int_{K_n^c} w(\cdot, y)\, \mu_n(dy)\right] \leqslant P_{K_n^c}\sigma$$

which decreases to zero as $n \longrightarrow \infty$, on $\{\sigma < \infty\}$ by Proposition 11 of § 1. The rest of the proof of (β) and (γ) are the same as before. To prove the new conclusion (δ) put $F = \{x : \sigma(x) < \infty \text{ and } s(x) < \infty\}$, and let $\mu(\partial D) = 0$. By (9), we have $L(x, \partial D) = 0$; by (5), we have $L_n(x, D) \longrightarrow L(x, D)$; both provided $x \in F$. Now for any measure ν define its part in D by

$$\nu^D(A) = \nu(D \cap A). \tag{14}$$

Then we have just shown that on F

$$U\mu_n^D \longrightarrow U\mu^D. \tag{15}$$

By (γ)

$$s = U\mu^D + U\mu^{E-D}; \tag{16}$$

hence on $\{s < \infty\}$, both terms on the right are excessive, and consequently

$$U\mu^D = \underline{U}\mu^D. \tag{17}$$

Since $\mu_n(Z) = 0$, we have by the definition of Z

$$U\mu_n^D = W\mu_n^D \qquad [= \textstyle\int w(x, y)\, \mu_n^D(dy)]. \tag{18}$$

Hence by the round property of w and the domination in condition (a), we have on F

$$P_D U\mu_n^D = U\mu_n^D. \tag{19}$$

Now F^c is polar by Proposition 8 of § 1, hence $P_D(x, F^c) = 0$ for every x. Therefore (15), valid on F, implies that

$$P_D U\mu_n^D \longrightarrow P_D U\mu^D \tag{20}$$

on F, by the domination in condition (a). It follows from (15), (19) and (20) that

$$P_D U\mu^D = U\mu^D \tag{21}$$

on F, hence by (17), using again $P_D(x, F^c) = 0$,

$$P_D \underline{U}\mu^D = \underline{U}\mu^D \tag{22}$$

everywhere by excessiveness. Thus for μ_D-a.e. y,

$$P_D \underline{u}(x, y) = \underline{u}(x, y) \tag{23}$$

first for ξ-a.e. x, then for all x. Let $\{D_n\}$ form a base of the topology such that $\mu(\partial D_n) = 0$. It follows from (23) that there exists N with $\mu(N) = 0$ such that if $y \notin N$, then for all n such that $y \in D_n$: $P_{D_n}\underline{u}(\cdot, y) = \underline{u}(\cdot, y)$; and therefore also for every G containing y, $P_G \underline{u}(\cdot, y) = \underline{u}(\cdot, y)$. For such a y the construction in Theorem 1 yields

$$\underline{u}(\cdot, y) = w(\cdot, y). \tag{24}$$

Since $\mu(N) = 0$, we have, using (γ):

$$s = U\mu = \underline{U}\mu = W\mu. \tag{25}$$

Since $u \geqslant \underline{u} \geqslant w$, (25) implies that $\mu(Z) = 0$. □

COROLLARY 4. — *The measure μ in Corollary 1 or 2 to Theorem 2 does not charge* Z.

By contrast, the method does not prove that the measure μ in Corollary 3 does not charge Z. This should be compared with (5) of § 1, where the measure μ is seen not to charge Z by the proof in [2] and the argument used in Theorem 2 to deduce (δ). Thus we have the anomaly of two representations of $P_K 1$ as potentials of measures both of which lack an essential feature. The next proposition clarifies the issue.

Recall that Hunt's Hypothesis (B) (see [3]) may be stated as follows: for every (open) G which contains (compact) K, we have

$$P_G P_K 1 = P_K 1. \tag{26}$$

There are several equivalent properties in terms of the sample function behavior of the process; see, e.g., [5]. It is known that the hypothesis is true under strong duality assumptions, see [5].

THEOREM 4. – *The following three propositions are equivalent.*

(a) Z *is polar;*

(b) *Hypothesis (B) is true;*

(c) $\forall K: P_K 1 = U\mu$ *where* $\lfloor\!\underline{\mu} \subset K$ *and* $\mu(Z) = 0$.

Proof. – (a) $\Longrightarrow$ (b): Suppose Z is polar and K be given. Let L be compact, $L \subset K \cap Z^c$. By Corollary 3 to Theorem 2, we have $P_L 1 = \underline{U}\mu$ where $\lfloor\!\underline{\mu} \subset L$. For each $y \in L$, we have $w(\cdot, y) = u(\cdot, y)$, hence also $= \underline{u}(\cdot, y)$. Thus we have $P_L 1 = W\mu$; and for any $G \supset K$

$$P_G P_L 1 = P_G W\mu = W\mu = P_L 1. \tag{27}$$

By Hunt's approximation theorem, $\exists L_n \subset K \cap Z^c$, $L_n \uparrow$, such that for each x, both P^x-a.s. and P^λ-a.s., where $\lambda(\cdot) = P_G(x, \cdot)$, we have $P_{L_n} 1 \uparrow P_{K \cap Z^c} 1$. The limit above is equal to $P_K 1$ because Z is polar. Taking such a sequence $\{L_n\}$ in (27), we obtain (26) by monotone convergence.

(b) $\Longrightarrow$ (c): Let $D_n \downarrow\downarrow K$; then by (b)

$$P_K 1 = P_{D_n} P_K 1 = U\mu_n , \tag{28}$$

where $\mu_n(Z) = 0$ and $\lfloor\!\underline{\mu_n} \subset \overline{D}_n$ by Corollaries 1 and 4 to Theorem 2. Hence by Theorem 2 (δ) there exists $\{\mu_{n_j}\}$ converging vaguely to μ so that (c) is true.

(c) $\Longrightarrow$ (a): Let $K \subset Z$; then (c) clearly implies that $P_K 1 \equiv 0$. Thus K is polar and so is Z.

We shall prove later that under the additional assumption that $\forall x: \underline{u}(x, x) = +\infty$, Z is indeed a polar set. Let us remark here that it is easy to show that Z is left-polar, hence semipolar. For this purpose we define

$$S_A = \inf\{t > 0: X_{t-} \in A\}, \quad P_A^* 1(x) = P^x\{S_A < \infty\}. \tag{29}$$

Then the method of proof in [2] yields

$$P_K^* 1 = U\nu, \quad \lfloor\!\underline{\nu} \subset K, \quad \nu(Z) = 0. \tag{30}$$

Hence the argument leading from (c) to (a) above shows that $P_Z^* 1 \equiv 0$.

4.

The main result of this section is that $U\mu$ uniquely determines μ provided that $\mu(Z) = 0$. This will be proved in a series of lemmas beginning with one due to Mokobodzki, which is essential. This is his result on excessivization valid for any discrete potential kernel; see [6] and [8]. The application to our case is made through standard techniques via resolvents, see [4]. Recall our convention in § 1 that all functions are positive measurable.

A "strong order" is defined as follows: $f \ll g$ iff $\exists$ an excessive function φ such that $f + \varphi = g$. For any measurable f (not necessarily positive), there exists a "least excessive majorant" f^* such that f^* is excessive, $f^* \geqslant f$, and for any excessive $\varphi \geqslant f$ we have $\varphi \geqslant f^*$. Mokobodzki's main result may be stated as follows:

$$\text{If} \quad f \ll g, \quad \text{then} \quad f^* \ll g^*. \tag{1}$$

We need also the following result, due to Mokobodzki (see another proof by Getoor in [7]).

(2) If φ is excessive and $\varphi \ll Ug$, then $\exists f$ such that $\varphi = Uf$.

Finally we need the following elementary uniqueness result.

$$\text{If} \quad Uf = Ug, \quad \text{then} \quad f = g \quad \xi\text{-a.e.} \tag{3}$$

This follows from a uniqueness theorem for additive functionals (see [1], p. 157), according to which the hypothesis in (3) implies that we have $u(x, y) f(y) \xi(dy) = u(x, y) g(y) \xi(dy)$ as measures for each x for which $Uf(x) < \infty$. Multiply both sides by $\dfrac{1}{u(x, y)}$ we obtain the conclusion in (3). For a general argument see [1].

Our application of these results is contained in the next lemma, which does not depend on the specific setting of this paper.

LEMMA 1. – *Let s be an excessive function which is finite ξ-a.e.; and let $Uf_n < \infty$ and*

$$Uf_n \uparrow s \tag{4}$$

everywhere. Suppose that we have

$$s = s_1 + s_2 \tag{5}$$

where s_1 *and* s_2 *are excessive. Then* $\exists g_n$ *and* h_n *such that*

$$f_n = g_n + h_n\,, \quad \xi\text{-a.e.}, \tag{6}$$

and

$$\mathrm{U}g_n \uparrow s_1\,, \quad \mathrm{U}h_n \longrightarrow s_2\,. \tag{7}$$

Proof. – Define

$$\varphi_n = (\mathrm{U}f_n - s_2)^+\,. \tag{8}$$

Then $\varphi_n \uparrow$ and

$$\mathrm{U}f_n = \varphi_n + (\mathrm{U}f_n \wedge s_2)\,. \tag{9}$$

Thus $\varphi_n \ll \mathrm{U}f_n$ and so by (1), $\varphi_n^* \ll \mathrm{U}f_n$. By (2), $\exists g_n$ and h_n such that $\varphi_n^* = \mathrm{U}g_n$ and

$$\mathrm{U}f_n = \mathrm{U}g_n + \mathrm{U}h_n\,. \tag{10}$$

This implies (6) by (3). Since $\varphi_n \leqslant s_1$ by (5) and (8), we have $\varphi_n^* \leqslant s_1$. Comparing (9) and (10), we see that $\mathrm{U}h_n \leqslant s_2$ because $\varphi_n \leqslant \mathrm{U}f_n$. Since $\varphi_n \uparrow$ so does $\varphi_n^* = \mathrm{U}g_n$. Let

$$\lim_n \uparrow \mathrm{U}g_n = \varphi\,; \quad \varliminf_n \mathrm{U}h_n = \psi\,. \tag{11}$$

Then φ is excessive and ψ is superaveraging. Letting $n \longrightarrow \infty$ in (10) and using (4) and (11), we see that

$$s = \varphi + \psi\,. \tag{12}$$

But $\varphi \leqslant s_1$ and $\psi \leqslant s_2$, hence in view of (5) we must have

$$\varphi = s_1\,, \quad \psi = s_2 \quad \text{on} \quad \{s < \infty\}\,. \tag{13}$$

Since $s < \infty$ ξ-a.e., this implies that $\underline{\psi} = s_2$, hence also $\psi = s_2$. Finally, $\varlimsup_n \mathrm{U}h_n \leqslant s_2$, hence $\lim_n \mathrm{U}h_n$ exists and $= s_2$. □

LEMMA 2. – *Let* s *be excessive,* $s \not\equiv \infty$, *then*

$$\mathrm{P_D}s = \mathrm{U}\mu \tag{14}$$

where $\mu(\mathrm{Z}) = 0$, $\underline{\lfloor \mu} \subset \overline{\mathrm{D}}$; *and this* μ *has the following splitting property. If* s_1 *and* s_2 *are excessive and*

$$\mathrm{P_D}s = s_1 + s_2\,, \tag{15}$$

then $\exists \mu_1$ *and* μ_2 *with* $\mu_1(\mathrm{Z}) = \mu_2(\mathrm{Z}) = 0$, *and*

$$\mu = \mu_1 + \mu_2\,, \quad s_1 = \mathrm{U}\mu_1\,, \quad s_2 = \mathrm{U}\mu_2\,. \tag{16}$$

Proof. – Except for the splitting property this has been stated in Corollaries 1 and 4 to Theorem 2. By Corollary 1 to Theorem 2, we have $\mathrm{U}f_n \uparrow \mathrm{P_D}\, s$ such that

$$f_n(y)\, \xi(dy) \longrightarrow \mu(dy)\,, \qquad \mathrm{P_D}\, s = \mathrm{U}\mu\,. \tag{17}$$

Hence by Lemma 1, $\exists\{g_n\}$ and $\{h_n\}$ satisfying (6) and (7). We can now apply Theorem 2 to $\{\mathrm{U}g_n\}$ and $\{\mathrm{U}h_n\}$ to obtain $\{n_j\}$ such that

$$\begin{aligned} g_{n_j}(y)\, \xi(dy) &\longrightarrow \mu_1(dy)\,, \quad s_1 = \mathrm{U}\mu_1\,, \quad \mu_1(\mathrm{Z}) = 0\,; \\ h_{n_j}(y)\, \xi(dy) &\longrightarrow \mu_2(dy)\,, \quad s_2 = \mathrm{U}\mu_2\,, \quad \mu_2(\mathrm{Z}) = 0\,. \end{aligned} \tag{18}$$

It is clear from (6), (17) and (18) that $\mu = \mu_1 + \mu_2$.

LEMMA 3. – *For any* y_0 *and* $x \notin \overline{\mathrm{D}}$, *we have*

$$\mathrm{P_D}\underline{u}(x\,, y_0) < \infty\,. \tag{19}$$

Proof. – By Corollary 1 to Theorem 2,

$$\mathrm{P_D}\underline{u}(x\,, y_0) = \int u(x\,, y)\, \mu(dy)\,, \quad \lfloor\underline{\mu} \subset \overline{\mathrm{D}}\,.$$

If $x \notin \overline{\mathrm{D}}$, then $\sup\limits_{y \in \mathrm{D}} u(x\,, y) < \infty$; from which (19) follows since μ is a Radon measure.

LEMMA 4. – *Suppose* $\mathrm{U}\mu = \mathrm{U}\nu \not\equiv \infty$, *where* $\mu(\mathrm{Z}) = \nu(\mathrm{Z}) = 0$. *If* $\lfloor\underline{\mu} \subset \mathrm{K}$, *then* $\lfloor\underline{\nu} \subset \mathrm{K}$.

Proof. – By Corollary to Proposition 8 in § 1, $\mathrm{U}\mu < \infty$ ξ-a.e. and both μ and ν are Radon measures. Let $\mathrm{F} = \{x \mid \mathrm{U}\mu(x) < \infty\}$, then $\xi(\mathrm{F}^c) = 0$. We have $\mathrm{U}\mu = \mathrm{W}\mu$, $\mathrm{U}\nu = \mathrm{W}\nu$. For any $\mathrm{D} \supset \mathrm{K}$, we have $\mathrm{P_D}\, \mathrm{W}\mu = \mathrm{W}\mu$ since $\lfloor\underline{\mu} \subset \mathrm{K}$; hence also

$$\mathrm{P_D}\, \mathrm{W}\nu = \mathrm{W}\nu\,. \tag{20}$$

Now if $x \in \mathrm{F}$, we can apply Fubini to infer from (20) that

$$\mathrm{P_D}\, w(x\,, y) = w(x\,, y) \tag{21}$$

for ν-a.e. y. Again by Fubini, $\exists\mathrm{N}$ with $\nu(\mathrm{N}) = 0$ such that if $y \notin \mathrm{N}$, then (21) holds for ξ-a.e. x, hence for all x because of excessiveness. If $\nu(\overline{\mathrm{D}}^c) > 0$, then $\exists y_0 \in (\overline{\mathrm{D}} \cup \mathrm{N} \cup \mathrm{Z})^c$ for which (21) holds with $x = y = y_0$. Since $y_0 \notin \mathrm{Z}$, $w(y_0\,, y_0) = \infty$, this would contradict Lemma 3. Thus $\lfloor\underline{\nu} \subset \overline{\mathrm{D}}$ and so $\lfloor\underline{\nu} \subset \mathrm{K}$ since D is arbitrary.

Z It is manifestly false that $W\mu = W\nu \not\equiv \infty$ and $\lfloor\mu \subset K$ implies $\lfloor\nu \subset K$.

THEOREM 5. – *Suppose that* $U\mu = U\nu \not\equiv \infty$ *where* $\mu(Z) = \nu(Z) = 0$; *then* $\mu \equiv \nu$.

Proof. – Suppose first that $\lfloor\mu \subset L$ where L is compact; then for any $D \supset L$ we have $U\mu = W\mu = P_D W\mu$. Hence by Lemma 2, $\exists$ a measure λ such that $\lambda(Z) = 0$,

$$W\mu = W\lambda \tag{22}$$

and moreover λ has the splitting property. We are going to show that $\mu \equiv \lambda$. The same argument then applies to ν by virtue of Lemma 4, and so $\nu \equiv \lambda \equiv \mu$.

Consider the Radon-Nikodym derivative $f = \dfrac{d\mu}{d(\mu + \lambda)}$ and $K \subset \left\{ f > \dfrac{1}{2} \right\}$. Recall the notation (14) in § 3. By Lemma 2, $\exists$ measure λ_1 such that $\lambda_1 \leqslant \lambda$ setwise and

$$W\mu^K = W\lambda_1 ; \tag{23}$$

by Lemma 4, $\lfloor\lambda_1 \subset K$. Were it possible that $(\mu + \lambda)(K) > 0$, then it would follow that

$$W\lambda^K \geqslant W\lambda_1^K = W\mu^K > \int_K w(x, y) \frac{1}{2} [\mu(dy) + \lambda(dy)] \tag{24}$$

and consequently by subtraction, on the set where $W\lambda < \infty$,

$$\int_K w(x, y) \lambda(dy) > \int_K w(x, y) \mu(dy) > \int_K w(x, y) \lambda(dy) . \tag{25}$$

This contradiction shows that $f \leqslant \dfrac{1}{2}$, $(\mu + \lambda)$-a.e.; which means $\mu \leqslant \lambda$. Together with (22) and $w(x, y) = u(x, y) > 0$ for $y \notin Z$, we conclude $\mu = \lambda$ as desired.

Coming to the general case of the theorem, we write

$$\begin{aligned} \varphi &= W\mu = W\mu^D + W\mu^{E-D} \\ P_D \varphi &= W\mu^D + P_D W\mu^{E-D} = W\mu^D + W\mu'_D \end{aligned} \tag{26}$$

where μ'_D is given by Corollaries 1 and 4 to Theorem 2, with $\lfloor\mu'_D \subset \overline{D}$. A similar expression holds when μ is replaced by ν in (26); and

so we obtain

$$W(\mu^{D} + \mu'_{D}) = W(\nu^{D} + \nu'_{D}). \tag{27}$$

Both measures here have support in $\overline{D}$, hence by what we have proved above,

$$\mu^{D} + \mu'_{D} \equiv \nu^{D} + \nu'_{D}. \tag{28}$$

Now on the set $F = \{x: W\mu(x) < \infty\}$, we have $W\mu^{E-D} \downarrow 0$ as $D \uparrow E$; a fortiori,

$$\lim_{D\uparrow E} W\mu'_{D} = \lim_{D\uparrow E} W\nu'_{D} = 0. \tag{29}$$

On the other hand, we obtain from (28) for each K

$$\int_{K} w(x, y)\, \mu^{D}(dy) + \int_{K} w(x, y)\, \mu'_{D}(dy) = \int_{K} w(x, y)\, \nu^{D}(dy) + \int_{K} w(x, y)\, \nu'_{D}(dy). \tag{30}$$

If $x \in F$, the second terms on both sides above converge to zero as $D \uparrow E$ by (29), whereas there is monotone convergence for the first terms. Therefore we have

$$\int_{K} w(x, y)\, \mu(dy) = \int_{K} w(x, y)\, \nu(dy),$$

and since K is arbitrary, the finite measures $w(x, y)\, \mu(dy)$ and $w(x, y)\, \nu(dy)$ coincide for $x \in F$. Fix such an x, and remember that w may be replaced by u which is strictly positive everywhere. We reach the final conclusion that $\mu \equiv \nu$. □

COROLLARY 5 to THEOREM 2. — *If $\mu_n(Z) = 0$ for all n, then conclusion (α) may be strengthened to read: μ_n converges vaguely to μ.*

This follows because all vague limits are the same by uniqueness.

We proceed to Riesz's decomposition. A function h is harmonic iff for every (compact) K

$$h = P_{K^c} h. \tag{31}$$

THEOREM 6. — *Let f be excessive, $\not\equiv \infty$. Then there exists a Radon measure μ with $\mu(Z) = 0$ and a harmonic function h such that*

$$f = U\mu + h. \tag{32}$$

If $f = U\mu_1 + h_1$ is another such representation (with $\mu_1(Z) = 0$), then $\mu \equiv \mu_1$ and $h \equiv h_1$.

Proof. – Let $K_n \uparrow E$ and $T_n = T_{K_n^c}$. Put

$$h = \lim_n P_{T_n} f \tag{33}$$

$$g = \begin{cases} f - h & \text{on } \{f < \infty\}, \\ \infty & \text{on } \{f = \infty\}. \end{cases} \tag{34}$$

Then we have

$$f = g + h. \tag{35}$$

It is clear from (33) that for each n,

$$h = P_{T_n} h \quad \text{on} \quad \{h < \infty\}; \tag{36}$$

h is superaveraging and $\{h \neq \underline{h}\} \subset \{h < \infty\} \subset \{f < \infty\}$ by (33) and the Lemma in § 1. Hence $\{h \neq \underline{h}\}$ is polar by Proposition 8 of § 1. Consequently $P_{T_n} h = P_{T_n} \underline{h}$, and so $\underline{h} = P_{T_n} \underline{h}$ except for a polar set and therefore everywhere. Thus $\underline{h}$ is harmonic.

Next we have from (35) and (36)

$$P_{T_n} f = P_{T_n} g + h \quad \text{on} \quad \{h < \infty\};$$

and so by (33),

$$h = \lim_n P_{T_n} g + h. \tag{37}$$

This shows that the limit in (37) is equal to zero on $\{h < \infty\}$, hence ξ-a.e. Assuming for a moment that g is superaveraging, then we have $\lim_n P_{T_n} \underline{g} = 0$ ξ-a.e., and this implies by standard arguments that $\underline{g}$ is a potential. Hence by Corollaries 1 and 4 of Theorem 2, $\underline{g} = U\mu$ with μ as asserted in the theorem. From (35) we have $f = \underline{g} + \underline{h} = U\mu + \underline{h}$ which is (32) except $\underline{h}$ is written there as h. The uniqueness is immediate by Theorem 5.

It remains to show that g is superaveraging. This is usually done via a result by Dynkin (see, e.g., [1], p. 273), but here is a shorter direct proof. Since h is the decreasing limit of excessive functions, we have $P_T h \leqslant P_S h$ if S and T are optional times such that $S \leqslant T$. In particular, we have by (31)

$$\forall t \geqslant 0: \ h = P_{T_n \wedge t} h, \quad \text{on} \quad \{h < \infty\}. \tag{38}$$

Now we have for each t

$$P_t g(x) \leqslant P_{T_n \wedge t} g(x) + E^x\{g(X_t); T_n \leqslant t\}. \tag{39}$$

Fix an x such that $f(x) < \infty$. Then the last term in (39) is bounded by $P^x\{f(X_t); T_n \leqslant t\}$ which converges to zero as $n \longrightarrow \infty$ since $P_t f(x) < \infty$. Furthermore, $P_{T_n \wedge t}(x, .)$ does not charge $\{f = \infty\}$ and so by (39) and the definition of g, we have

$$P_t g(x) \leqslant \varliminf_n P_{T_n \wedge t} g(x) = \varliminf_n [P_{T_n \wedge t} f(x) - P_{T_n \wedge t} h(x)]$$
$$= \varliminf_n P_{T_n \wedge t} f(x) - h(x) \leqslant f(x) - h(x) \leqslant g(x).$$

Thus $P_t g \leqslant g$ on $\{f < \infty\}$; hence everywhere since $g = \infty$ on $\{f = \infty\}$. □

PROPOSITION 13. – *For each y except possibly a polar set, we have*

$$\lim_{K \uparrow E} P_{K^c} u(x, y) = 0, \quad \textit{for} \quad x \neq y; \tag{40}$$

$$\lim_{t \to \infty} P_t u(x, y) = 0, \quad \textit{for} \quad x \neq y. \tag{41}$$

Proof. – Applying the Lemma in § 1, with $f = u(\cdot, y)$ and using the notation there, we have for fixed y

$$\lim_{n \to \infty} P_{T_n} u(x, y) = g(x, y)$$

where $g(x, y) = \underline{g}(x, y)$ if $x \neq y$. By Proposition 9 of § 1, either $\underline{g}(\cdot, y) \equiv 0$, or $\underline{g}(\cdot, y) > 0$ everywhere. Let $K \subset \{y : \underline{g}(\cdot, y) \not\equiv 0\}$. By Corollary 3 of Theorem 2, $\exists \mu$ with $\lfloor\mu \subset K$ such that

$$P_K 1(x) = \int u(x, y)\, \mu(dy) \qquad \text{if} \quad x \notin K.$$

Hence by transience and dominated convergence,

$$0 = \lim_n P_{T_n} P_K 1(x) = \int g(x, y)\, \mu(dy).$$

But if $y \in K$, $g(x, y) \geqslant \underline{g}(x, y) > 0$. Hence $\mu \equiv 0$ and so K is polar by the cited Corollary 3. This proves (40), and (41) is similar.

An excessive function s is called "purely excessive" iff

$$\lim_{t \to \infty} P_t s = 0 \qquad \xi\text{-a.e.} \tag{42}$$

A "pure potential" is a potential which is purely excessive. A result analogous to Proposition 11 of § 1 shows that (42) implies actually the limit there is zero on $\{s < \infty\}$, hence except possibly a polar set.

The following remark is important. If s is purely excessive, then we have everywhere

$$Uf_n \uparrow s\,, \quad \text{where} \quad f_n = n(s - P_{1/n}s)\,. \tag{43}$$

To see this, recall that standard arguments show that Uf_n is increasing and converges to s on the set where (42) holds. Hence the limit is an excessive function which is equal to s ξ-a.e., therefore it coincides with s.

5.

Hunt's Hypothesis (B), which is equivalent to

$$Z \text{ is a polar set} \tag{1}$$

by Theorem 4, will now be proved under the additional assumption below:

$$\forall y : \underline{u}(y\,, y) = +\infty\,. \tag{2}$$

This is satisfied if $\underline{u} = u$, namely if $u(\cdot\,, y)$ is excessive for each y. The latter condition is in turn satisfied if $u(\cdot\,, y)$ is lower semi-continuous since it is superaveraging by Fatou's lemma. At a crucial point we need also the assumption

$$\xi \text{ is an excessive measure;} \tag{3}$$

namely $\xi \geqslant \xi P_t$ for every $t \geqslant 0$. This assumption is usually made for a reference measure.

Define the Borel set

$$Q = \{y : \underline{u}(\cdot\,, y) \quad \text{is a pure potential}\}\,. \tag{4}$$

According to Proposition 13 in § 4, Q^c is a polar set. We shall prove that $Q \subset Z^c$; then $Z \subset Q^c$ so that (1) is true.

Let $y \in Q$. Then for each $t \geqslant 0$, $P_t\underline{u}(\cdot\,, y)$ is a potential. Hence by Corollaries 1 and 4 of Theorem 2, there exists a Radon measure, to be denoted by $\hat{P}_t(\cdot\,, y)$, which does not charge Z, such that

$$P_t\underline{u}(\cdot\,, y) = U\hat{P}_t(\cdot\,, y) = \int u(\cdot\,, z)\,\hat{P}_t(dz\,, y)\,. \tag{5}$$

Here we have adopted the left-handed notation appropriate for the dual symbolism. Note that we can replace u by $\underline{u}$ in the last term

above, by Proposition 4 of § 1. It follows from (5) that if $y \in Q$, then $\hat{P}_t(\cdot, y)$ does not charge Q^c because the first member of (5) is a pure potential.

THEOREM 7. – *$\{\hat{P}_t, t \geq 0\}$ is a semigroup of kernels on $Q \times Q$. We have for each $y \in Q$*

$$\hat{P}_t(\cdot, y) \longrightarrow \hat{P}_0(\cdot, y) \quad \textit{vaguely.} \tag{6}$$

If we define a kernel $\hat{U}$ on $Q \times Q$ as follows:

$$\hat{U}(dx, y) = \xi(dx)\, u(x, y), \tag{7}$$

then for any f (≥ 0) for which $f\hat{U}(y) < \infty$ we have

$$f\hat{U}(y) = \int_0^\infty \hat{P}_t(f, y)\, dt. \tag{8}$$

Proof. – The key is the uniqueness Theorem 5. We have for $w \in Q$, $t \geq 0$ and $s \geq 0$, since $\hat{P}_s(\cdot, w)$ is concentrated on Q, by (5):

$$\begin{aligned}\int [u(x, z)\, \hat{P}_t(dz, y)]\, \hat{P}_s(dy, w) &= \int P_t \underline{u}(x, y)\, \hat{P}_s(dy, w)\\ &= \int P_t(x, d\eta) \left[\int \underline{u}(\eta, y)\, \hat{P}_s(dy, w)\right]\\ &= \int P_t(x, d\eta)\, P_s \underline{u}(\eta, w) = P_{t+s}\underline{u}(x, w)\\ &= \int u(x, z)\, \hat{P}_{t+s}(dz, w).\end{aligned}$$

Hence

$$\int \hat{P}_t(dz, y)\, \hat{P}_s(dy, w) = \hat{P}_{t+s}(dz, w)$$

which establishes the semigroup property. Let $t_n \downarrow 0$ and apply Theorem 2 to the sequence $U\hat{P}_{t_n} = P_{t_n}\underline{u}$, with $\sigma = s = \underline{u}(\cdot, y)$ and under condition (c_2), we obtain $\underline{u}(\cdot, y) = U\mu$ where μ is a vague limit of $\hat{P}_{t_n}$, with $\mu(Z) = 0$. Hence $\mu = \hat{P}_0$ by Theorem 5, and (6) follows.

Next we have by Fubini,

$$\begin{aligned}U[\underline{u}(\cdot, y) - P_t \underline{u}(\cdot, y)] &= \int_0^t P_s \underline{u}(\cdot, y)\, ds \qquad (9)\\ &= \int_0^t ds \int u(\cdot, z)\, \hat{P}_s(dz, y) = \int u(\cdot, z) \int_0^t \hat{P}_s(dz, y)\, ds.\end{aligned}$$

The first term above is equal to

$$\int u(\cdot, z)\, [\underline{u}(z, y) - P_t \underline{u}(z, y)]\, \xi(dz);$$

hence by the uniqueness theorem we have

$$[\underline{u}(z, y) - P_t\underline{u}(z, y)]\,\xi(dz) = \int_0^t \hat{P}_s(dz, y)\,ds. \tag{10}$$

For any f $(\geqslant 0)$ such that

$$\int \xi(dz)\, f(z)\, u(z, y) < \infty \tag{11}$$

we have

$$\int \xi(dz)\, f(z)\, [\underline{u}(z, y) - P_t\underline{u}(z, y)] = \int_0^t f\hat{P}_s(y)\,ds,$$

and so letting $t \longrightarrow \infty$

$$\int \xi(dz)\, f(z)\, \underline{u}(z, y) = \int_0^\infty f\hat{P}_s(y)\,ds. \tag{12}$$

Note that $\forall y: \underline{u}(z, y) = u(z, y)$ for ξ-a.e. z, hence it is immaterial whether u or $\underline{u}$ is written in (11) or (12).

Remark. – It can be shown that if y is not in a certain polar set, then (11) holds for any bounded f with compact support.

Theorem 7 requires an essential complement which is stated separately to stress the point. We need first a lemma, the only place where the excessiveness of ξ is used. We write $\xi(f)$ for $\int f(x)\,\xi(dx)$.

LEMMA 5. – *If*

$$Uf < \infty \quad \textit{and} \quad Uf \leqslant \varliminf_n Ug_n \quad \xi\text{-a.e.}, \tag{13}$$

then

$$\xi(f) \leqslant \varliminf_n \xi(g_n).$$

Proof. – We prove first that if $Uf < \infty$ and $Uf \leqslant Ug$, ξ-a.e., then $\xi(f) \leqslant \xi(g)$. For this purpose we may suppose $\xi(g) < \infty$, hence $\xi(U^\lambda g) < \infty$ for $\lambda > 0$ because $\lambda\xi U^\lambda \leqslant \xi$. Write $P_t^\lambda = e^{-\lambda t}P_t$; then we have for any fixed $\lambda > 0$:

$$\frac{1}{t}\,(\xi - \xi P_t^\lambda)\,(U^\lambda g) = \frac{1}{t}\int_0^t (\xi P_s^\lambda)\,(g)\,ds \uparrow \xi(g) \tag{14}$$

as $t \downarrow 0$. Hence if

$$U^\lambda f \leqslant U^\lambda g, \tag{15}$$

then $\xi(f) \leqslant \xi(g)$. [We learned this argument from M.J. Sharpe.] Unfortunately (15) is not part of our hypothesis; whereas (14) need not hold for $\lambda = 0$. The remedy is as follows. Let $0 < a < 1$ and

put for $n \geqslant 1$:

$$A_n = \{x \in E \mid aU^{1/n} f(x) \leqslant U^{1/n} g(x)\}.$$

Since $U^{1/n}$ increases to U as $n \longrightarrow \infty$, our hypotheses imply that $\xi(E - \liminf_n A_n) = 0$. Now we have on A_n

$$aU^{1/n}(f 1_{A_n}) \leqslant U^{1/n} f \leqslant U^{1/n} g ;$$

hence the inequality holds everywhere in E by the domination principle for $U^{1/n}$ (see [4; p. 245]). Therefore the argument above with $\lambda = \dfrac{1}{n}$ yields $a\xi(f 1_{A_n}) \leqslant \xi(g)$. Letting $n \longrightarrow \infty$, then $a \uparrow 1$, we obtain $\xi(f) \leqslant \xi(g)$.

Now suppose that (13) is true. For $0 < a < 1$ put

$$B_m = \{x \in E \mid aUf(x) < \inf_{n \geqslant m} U(g_n)\}.$$

Then $B_m \uparrow$ and $\xi(E - \bigcup_m B_m) = 0$. We have for each $n \geqslant m$: $aU(f 1_{B_m}) \leqslant U(g_n)$ on B_m; hence the inequality holds everywhere by the domination principle for U. It follows from the first part of the proof that $a\xi(f 1_{B_m}) \leqslant \xi(g_n)$. Letting $n \longrightarrow \infty$, we infer that $a\xi(f 1_{B_m}) \leqslant \underline{\lim}_{n \to \infty} \xi(g_n)$. Letting $m \longrightarrow \infty$, then $a \uparrow 1$, we obtain the conclusion of the lemma.

THEOREM 8. – $\{\hat{P}_t, t \geqslant 0\}$ *on* $Q \times Q$ *is a submarkovian semi-group, i.e.,* $\forall y \in Q : \hat{P}_t(Q, y) \leqslant 1$.

Proof. – Fix $y \in Q$, $t \geqslant 0$, and put for $\delta > 0$

$$f_\delta(x) = \frac{1}{\delta} [P_t \underline{u}(x, y) - P_{t+\delta} \underline{u}(x, y)].$$

Since $P_t \underline{u}(\cdot, y)$ is purely excessive, we have

$$Uf_\delta(x) \leqslant P_t \underline{u}(x, y) \tag{16}$$

$$\lim_{\delta \downarrow 0} \uparrow Uf_\delta(x) = P_t \underline{u}(x, y) \tag{17}$$

for all x, by (43) of § 4. Hence by Theorems 2 and 5 we have

$$f_{1/n}(z)\, \xi(dz) \longrightarrow \hat{P}_t(dz, y) \quad \text{vaguely.} \tag{18}$$

Let $D_n \downarrow\downarrow \{y\}$, and ($\xi$ charges each open set)

$$g_n(z) = \frac{1_{D_n}(z)}{\xi(D_n)} \tag{19}$$

so that

$$\forall n: \ \xi(g_n) = 1 . \tag{20}$$

For each x, $u(x, \cdot)$ is lower semicontinuous; since $g_n(z)\,\xi(dz)$ converges vaguely to the unit mass at y, we have

$$u(x, y) \leqslant \varliminf_n \mathrm{U}g_n(x) . \tag{21}$$

Therefore we have in conjunction with (16), for each $\delta > 0$,

$$\mathrm{U}f_\delta \leqslant u(\cdot, y) \leqslant \varliminf_n \mathrm{U}g_n . \tag{22}$$

It follows by Lemma 5 and (20) that

$$\xi(f_\delta) \leqslant 1 \tag{23}$$

for every δ, and consequently by (18) that $\hat{\mathrm{P}}_t(\mathrm{Q}, y) \leqslant 1$. □

A function s defined on Q is called "co-superaveraging" iff $s \geqslant s\hat{\mathrm{P}}_t$ for every $t > 0$ and is "co-excessive" iff in addition $s = \lim\limits_{t \downarrow\downarrow 0} s\hat{\mathrm{P}}_t$. We cannot yet define a "co-potential", but we can define a co-excessive s to be "purely co-excessive" iff $\lim\limits_{t\to\infty} s\hat{\mathrm{P}}_t = 0$, ξ-a.e.

The following lemma is the co-version of a remark at the end of § 4, and is spelled out here because of its importance in the proof of Theorem 10 below.

LEMMA 6. – *If φ is purely co-excessive, then*

$$\psi_n \hat{\mathrm{U}} \uparrow \varphi , \quad \textit{where} \quad \psi_n = n(\varphi - \varphi\,\hat{\mathrm{P}}_{1/n}) \tag{24}$$

everywhere.

Proof. – Just as in (43) of § 4, $\psi_n \hat{\mathrm{U}} \uparrow \widetilde{\varphi}$, where $\widetilde{\varphi}$ is co-excessive and $\widetilde{\varphi} = \varphi$ ξ-a.e. Now we see from (10) that the measure $\int_0^t \hat{\mathrm{P}}_s(\cdot, y)\,ds$, $y \in \mathrm{Q}$, is absolutely continuous with respect to ξ. Hence

$$\frac{1}{t} \int_0^t \widetilde{\varphi}\hat{\mathrm{P}}_s ds = \frac{1}{t} \int_0^t \varphi\hat{\mathrm{P}}_s ds . \tag{25}$$

Letting $t \downarrow 0$, we obtain $\widetilde{\varphi} = \varphi$. □

THEOREM 9. – *For each x, the function $\underline{u}(x, \cdot)$ on Q is purely co-excessive; so is*

$$\varphi_x(\cdot) = 1 - e^{-\underline{u}(x,\cdot)} . \tag{26}$$

Proof. – Using Proposition 4 of § 1, we see that (5) may be written as

$$\underline{u}_x \hat{P}_t(y) = P_t \underline{u}_x(y) \tag{27}$$

where $\underline{u}_x(y) = \underline{u}(x, y)$. This shows at once that u_x is purely co-excessive. Let $\Phi(\theta) = 1 - e^{-\theta}$, $\theta \geq 0$; then Φ is concave and $\varphi_x = \Phi \circ u_x$. Hence by Jensen's inequality

$$\varphi_x \geq \Phi \circ u_x \hat{P}_t \geq (\Phi \circ u_x) \hat{P}_t = \varphi_x \hat{P}_t, \tag{28}$$

namely φ_x is co-superaveraging. Now u_x is lower semicontinuous, so are successively: $P_t u_x$, $\underline{u}_x$ and φ_x. Therefore φ_x is co-excessive. It is purely so by (28), because u_x is purely co-excessive and $\Phi(0) = 0$. □

THEOREM 10. – $Q \subset Z^c$.

Proof. – Fix a $y \in Q$, and put $\varphi(z) = 1 - e^{-\underline{u}(y,z)}$. By Theorem 9, φ is purely co-excessive; hence by Lemma 6, $\exists \psi_n$ such that

$$\forall z: \ \psi_n \hat{U}(z) \uparrow \varphi(z). \tag{29}$$

Now by integrating the fundamental representation formula (5) with $t = 0$, we have

$$\int \xi(dx)\, \psi_n(x)\, \underline{u}(x, y) = \int \left[\int \xi(dx)\, \psi_n(x)\, \underline{u}(x, z) \right] \hat{P}_0(dz, y),$$

namely $\psi_n \hat{U}(y) = \int \psi_n \hat{U}(z)\, \hat{P}_0(dz, y)$. Letting $n \longrightarrow \infty$ and using (29), we obtain

$$\varphi(y) = \int \varphi(z)\, \hat{P}_0(dz, y). \tag{30}$$

Observe that $\varphi(y) = 1$ by our new assumption (2), and $\varphi(z) < 1$ for $z \neq y$ by our old assumption. Hence (20) together with $\hat{P}_0(Q, y) \leq 1$ (Theorem 8) forces

$$\hat{P}_0(\{y\}, y) = 1. \tag{31}$$

But $\hat{P}_0(Z, y) = 0$ by Theorem 2, as recalled before (5). Therefore $y \in Z^c$. □

Finally, we give a generalization of Theorem 10 by weakening the condition (2) as follows:

(32) the family of excessive functions $\underline{u}(\cdot, y)$ on E, indexed by $y \in E$, are all distinct.

Such a condition is meaningfull in the general theory of excessive functions. It is implied by (2) because $\underline{u}(y, y') < \infty = \underline{u}(y, y)$ if $y' \neq y$. Our basic assumption (iib) in § 1 is thus thrown into relief.

THEOREM 11. – *Under our basic assumptions (i) and (ii) stated at the beginning of* § *1, with* ξ *an excessive measure, if (32) is true, then Hunt's Hypothesis (B) is true.*

Proof. – Fix $y \in \mathrm{Q}$ and write for brevity's sake μ for the measure $\hat{\mathrm{P}}_0(\cdot, y)$. As an obvious generalization of (30), we have for each $\lambda > 0$:

$$1 - e^{-\lambda \underline{u}(x,y)} = \int [1 - e^{-\lambda \underline{u}(x,z)}]\, \mu(dz). \tag{33}$$

It follows from Proposition 9 of § 1 that $\underline{u}(x, y) > 0$ for all (x, y). Hence letting $\lambda \uparrow \infty$ above, we obtain $\mu(\mathrm{E}) = 1$. Recall that μ is concentrated on Z^c as well as on Q. Next we have by (5) with $t = 0$:

$$\underline{u}(x, y) = \int \underline{u}(x, z)\, \mu(dz). \tag{34}$$

If we put $\lambda = 1$ in (33), the resulting equation may be written with our previous notation Φ as follows:

$$\Phi(\underline{u}(x, y)) = \int \Phi(\underline{u}(x, z))\, \mu(dz). \tag{35}$$

Now Φ is strictly concave, whereas we have the equality case of Jensen's inequality in (35). This forces the measure μ to concentrate on the set of z where the integrand $u(x, z)$ in (34) takes a constant value, which must then be $\underline{u}(x, y)$ because $\mu(\mathrm{E}) = 1$. [We were unable to unearth a reference to the required proposition in the case of a general probability measure μ, but Michael Steele was kind enough to supply an elegant short proof on request.] Namely, if we put $\mathrm{B}_x = \{z \in \mathrm{Z}^c \,|\, \underline{u}(x, z) = \underline{u}(x, y)\}$, then $\mu(\mathrm{B}_x) = 1$ for each $x \in \mathrm{E}$. Now put also $\mathrm{C}_x = \{z \in \mathrm{E} \,|\, u(x, z) = \underline{u}(x, y)\}$. Since $z \longrightarrow u(x, z)$ is extended continuous, C_x is a closed set. We have $\mathrm{B}_x = \mathrm{C}_x \cap \mathrm{Z}^c$, because if $z \in \mathrm{Z}^c$, then $\underline{u}(x, z) = u(x, z)$ for all $x \in \mathrm{E}$. Let $\mathrm{B} = \bigcap_{x \in \mathrm{E}} \mathrm{B}_x$, $\mathrm{C} = \bigcap_{x \in \mathrm{E}} \mathrm{C}_x$. Since $\mathrm{C}_x \supset \mathrm{B}_x$, we have $\mu(\mathrm{C}_x) = 1$. Since C is closed, it follows that $\mu(\mathrm{C}) = 1$ a cute little exercise in measure theory. Therefore, $\mathrm{B} = \mathrm{C} \cap \mathrm{Z}^c$ belongs to $\mathfrak{B}$ and $\mu(\mathrm{B}) = 1$. Thus B is not empty. Let $y' \in \mathrm{B}$,

then we have by the definition of B: $\forall x \in E$: $\underline{u}(x, y') = \underline{u}(x, y)$. Our new condition (32) entails $y' = y$. Thus $y \in Z^c$ as we concluded at the end of the proof of Theorem 10. [In fact, $B = \{y\}$ and (31) follows.]

In closing, let us remark that the preceding proof of Hypothesis (B) is "perilously close", as Hunt would have said. More than one attempt was made to simplify it, but the efforts failed on rather delicate details. Hunt said [3, p. 81], "I have not found simple and general conditions on the transition measures to ensure the truth of Hypothesis (B)." it is implied by the usual duality assumptions, see [5]. It would be extremely interesting to know whether a simpler proof exists in the setting of this paper.

BIBLIOGRAPHY

[1] R.M. Blumenthal and R.K. Getoor, Markov Processes and Potential Theory, Academic Press, 1968.

[2] K.L. Chung, Probabilistic approach in potential theory to the equilibrium problem, *Ann. Inst. Fourier*, 23, 3 (1973), 313-322.

[3] G.A. Hunt, Markoff processes and potentials I, *Illinois J. Math.*, 1 (1957), 43-93.

[4] P.A. Meyer, Probabilités et potentiel, Hermann, 1966.

[5] P.A. Meyer, Processus de Markov : la frontière de Martin, *Lecture Notes in Mathematics* No. 77, Springer-Verlag, (1968).

[6] P.A. Meyer, Deux petits résultats de théorie du potentiel, Séminaire de Probabilités V, *Lecture Notes in Mathematics* No. 191, Springer-Verlag (1971), 211-212.

[7] P.A. Meyer, Le Retournement du temps, d'après Chung et Walsh, Séminaire de Probabilités V, *Lecture Notes in Mathematics* No. 191, Springer-Verlag (1971), 213-236.

[8] G. Mokobodzki, Densité relative de deux potentiels comparables, Séminaire de Probabilités IV, *Lecture Notes in Mathematics* No. 124, Springer-Verlag (1970), 170-194.

[9] K.M. RAO, Excessive functions as potentials of measures, *J. London Math. Soc.*, 16 (1977), 165-171.

Manuscrit reçu le 26 décembre 1979.

K.L. CHUNG,
Stanford University
Department of Mathematics
Stanford, Ca. 94305 (USA).

K. MURALI RAO,
Mathematics Institute
Aarhus University
Aarhus (Danemark).

université scientifique et médicale
de grenoble

annales de l'institut fourier

ERRATUM

"A NEW SETTING FOR POTENTIAL THEORY (part 1)"

Article paru dans le tome 30 (1980), fascicule 3, pp. 167-198

Mémoire de K. L. CHUNG & K. MURALI RAO

On p. 181, (26) should read as follows :

$$(26) \qquad P_G P_K = P_K .$$

Proof that (a) ⇒ (b) should be revised as follows.

Suppose Z is polar and K be given. Let L be compact, $L \subset K \cap Z^c$. By Proposition 1 of § 1, there exists h such that $h > 0$ everywhere and $Uh \leq 1$. Let $s = P_K Uh$, then $s = \lim_n \downarrow P_{D_n} Uh$ where $D_n \downarrow\downarrow K$. By Corollaries 1 and 4 of Theorem 2 (continued), we have

$$P_{D_n} Uh = U\mu_n \quad , \quad \underline{\mu}_n \subset \overline{D}_n \quad , \quad \mu_n(Z) = 0 .$$

Hence $s = \lim_n U\mu_n$, and $U\mu_n \leq P_{D_1} Uh < \infty$ for all n. Apply Theorem 2 (continued) under (c_1) to obtain $\{\mu_{n_j}\}$ converging vaguely to μ, such that $s = U\mu$ and $\mu(Z) = 0$, the last assertion by (6) of Theorem 2. We have $\underline{\mu} \subset K$ by vague convergence. Thus

$$s = U\mu \quad , \quad \underline{\mu} \subset K \quad , \quad \mu(Z) = 0 ,$$

and therefore $s = W\mu$. For any (open) $G \supset K$, we have then

$$P_G s = P_G W\mu = W\mu = s$$

where the second equation is due to the round property of w and the fact μ is supported by $K \subset G$. Thus by the argument on p. 70 of [5] :

- 2 -

$$0 = P_K Uh - P_G P_K Uh \geq E^{\cdot}\left\{\int_{T_K}^{T_G + T_K \circ \theta_{T_G}} h(X_t)dt \; ; \; T_G = T_K \; ; \; X(T_G) \in K \setminus K^r\right\}$$

which implies that

$$\forall x : \quad P^x\{T_G = T_K \; ; \; X(T_G) \in K \setminus K^r\} = 0 \; .$$

This implies easily that for any $f \in b\mathcal{E}_+$:

$$P_G P_K f = P_K f$$

which is (26).

N.B. The mistake was to suppose that $P_G P_K 1 = P_K 1$ implies $P_G P_K = P_K$. This was partly caused by a statement on p. 71 of [5] which apparently asserts that $P_G P_K \leq P_K$ in general. Dellacherie gave a trivial counter-example to the last assertion, which is left as an exercise.

First display on p. 168 should read :

$$\lim_{t \to \infty} P^x\{T_K \circ \theta_t < \infty\} = 0 \; .$$

Reprinted from *Seminar on Stochastic Processes, 1981*, Birkhäuser, Boston, 1981, pp. 1–29.
With kind permission from Springer Science and Business Media.

FEYNMAN-KAC FUNCTIONAL AND THE SCHRÖDINGER EQUATION*

by

K.L. CHUNG and K.M. RAO

The Feynman-Kac formula and its connections with classical analysis were inititated in [3]. Recently there has been a revival of interest in the associated probabilistic methods, particularly in applications to quantum physics as treated in [7]. Oddly enough the inherent potential theory has not been developed from this point of view. A search into the literature after this work was under way uncovered only one paper by Khas'minskii [4] which dealt with some relevant problems. But there the function q is assumed to be nonnegative and the methods used do not apparently apply to the general case; see the remarks after Corollary 2 to Theorem 2.2 below.** The case of q taking both signs is appealing as it involves oscillatory rather than absolute convergence problems. Intuitively, the Brownian motion must make intricate cancellations along its paths to yield up any determinable averages. In this respect Theorem 1.2 is a decisive result whose significance has yet to be explored. Next we solve the boundary value problem for the Schrödinger equation $(\Delta+2q)\varphi = 0$. In fact, for a

*Research supported in part by a NSF Grant MCS-8001540.

**The case $q \le 0$ is "trivial" in the context of this paper. For this case in a more general setting see [9, Chapter 13].

positive continuous boundary function f, a solution is obtained in the explicit formula given in (2) of §1 below, provided that this quantity is finite (at least at one point x in D). Thus the Feynman-Kac formula supplies the natural Green's operator for the problem. For a domain with finite measure, the result is the best possible as it includes the already classical solution of the Dirichlet problem by probability methods. Other results are valid for an arbitrary domain and it seems that some of them are proved here under less stringent conditions than usually given in non-probabilistic treatments. For instance, no condition on the smoothness of the boundary is assumed beyond that of regularity in the sense of the Dirichlet problem, and the basic results hold without this regularity. Of course, the Schrödinger equation is a case of elliptic partial differential equations on which there exists a huge literature, but we make no recourse to the latter theory. Comparisons between the methods should prove worthwhile and will be discussed in a separate publication.

It is well known that the Schrödinger equation differs essentially from the Laplace equation in that a condition on the size of the domain is necessary to guarantee the uniqueness of solution. In our context it is evident at the outset that the key to this is the quantity $u_D(x) = E^x\{ \exp(\int_0^{\tau_D} q(x(t))\, dt) \}$, the finiteness of which lies at the base of the probabilistic considerations. As natural as it is from our point of view, this quantity does not lend itself easily to non-probabilistic analysis. The identification in the simplest case (see the remark after Lemma D) as a particular solution of the equation is one of those amusing twists not uncommon in other theories when dealing with an object which has really a simple probabilistic existence.

The one-dimensional case of this investigation has appeared in [1] though the orientation is somewhat different there. A summary of

the present results has been announced in [2].

1. Harnack inequality; global bound; boundary limit

Let $\{ X(t),\ t \geq 0 \}$ be the Brownian motion process in R^d, $d \geq 1$; with all paths continuous. The transition semigroup is $\{ P_t,\ t \geq 0 \}$ and $\mathcal{F}_t$ is the σ-field generated by $\{ X_s,\ 0 \leq s \leq t \}$ and augmented in the usual way. The qualifying phrase "almost surely" (a.s.) will be omitted when readily understood. A "set" is always a Borel set and a "function" is always a Borel measurable function. The class of bounded functions will be denoted by $b\mathcal{B}$; if its domain is A this is indicated by $b\mathcal{B}(A)$. Similarly for other classes of functions to be used later. The sup-norm of $f \in b\mathcal{B}$ is denoted by $\|f\|$; restricted to A it is denoted by $\|f\|_A$. P^x and E^x denote the probability and expectation for the process starting at x.

For any set B we put

$$\tau(B) = \tau_B = \inf\{ t > 0 \mid X(t) \notin B \} \quad ;$$

namely the first exit time from B, with the usual convention that $\inf \phi = \infty$. Let $q \in b\mathcal{B}$; as an abbreviation we put

$$(1) \qquad e_q(t) = \exp\{ \int_0^t q(X(s))\, ds \} \quad ;$$

when q is fixed it will be omitted from the notation. A domain in R^d is an open connected set; its boundary is $\partial D = \bar{D} \cap \overline{D^c}$, where $\bar{D}$ is the closure and D^c the complement of D. For $f \geq 0$ on ∂D we put for $x \in \bar{D}$:

$$(2) \qquad u(q,f;x) = E^x\{ e_q(\tau_D)\, f(X(\tau_D));\ \tau_D < \infty \}.$$

The following result is a case of Harnack's inequality, on which there is a considerable literature for elliptic partial differential equations.

Theorem 1.1. Let D be a domain and K a compact subset of D. There exists a constant $A > 0$ which depends only on D, K and Q, such that for any q with $\|q\| \le Q$ and $f \ge 0$ such that $u(q,f;\cdot)\not\equiv\infty$ in D, we have for any two points x_1 and x_2 in K:

$$(3) \qquad A^{-1}\, u(q,f;x_2) \le u(q,f;x_1) \le A\, u(q,f;x_2) \quad .$$

Proof. We write $u(x)$ for $u(q,f;x)$. By hypothesis there exists $x_0 \in D$ such that $u(x_0) < \infty$. We may suppose $x_0 \in K$ by enlarging K. For any $r > 0$ define

$$T(r) = \inf\{\, t > 0 \mid \rho(X(t), X(0)) \ge r \,\}$$

where ρ denotes the Euclidean distance. It is well known (cf. Lemma A below) that there exists $\delta > 0$ (which depends only on Q and the dimension d) such that for all $x \in R^d$:

$$(4) \qquad \frac{1}{2} \le E^x\{\, \exp(-QT(2\delta)) \,\}; \qquad E^x\{\, \exp(QT(2\delta)) \,\} \le 2.$$

In fact, the two expectations in (4) do not depend on x by the spatial homogeneity of the process. Now put

$$(5) \qquad 2r = \rho(K, \partial D) \wedge 2\delta.$$

Then for any $s < 2r$ we have, by the strong Markov property, since $T(s) < \tau_D$ under P^{x_0}:

$$(6) \qquad \infty > u(x_0) = E^{x_0}\{ e(T(s))\, u(X(T(s))) \}$$

$$\geq E^{x_0}\{ \exp(-QT(s))\, u(X(T(s)) \} .$$

The isotropic property of the Brownian motion implies that the random variables $T(s)$ and $X(T(s))$ are stochastically independent for each s. Hence we obtain from (6) and the first inequality in (4):

$$(7) \qquad u(x_0) \geq \frac{1}{2} E^{x_0}\{ u(X(T(s))) \} .$$

The expectation on the right side above is the area average of the values of u on the boundary of $B(x_0,s)$. Hence we obtain by integrating with respect to the radius:

$$(8) \qquad a_d \int_0^{2r} E^{x_0}\{ u(X(T(s))) \}\, s^{d-1}\, ds = \int_{B(x_0,2r)} u(y)\, dy,$$

where $a_d\, s^{d-1}$ is the area of $\partial B(x_0,s)$. It follows from (7) and (8) that

$$(9) \qquad u(x_0) \geq \frac{1}{2V(2r)} \int_{B(x_0,2r)} u(y)\, dy,$$

where $V(2r)$ is the volume of $B(x_0,2r)$. [The terms "area" and "volume" used above have their obvious meanings in dimension $d = 1$ or 2.]

Next, let $x \in B(x_0,r)$ so that $\rho(x,\partial D) \geq r$ by (5). We have for $0 < s < r$:

$$(10) \qquad u(x) = E^{x}\{ e(T(s))\, u(X(T(s))) \} \leq E^{x}\{ \exp(QT(s))\, u(X(T(s))) \}$$

$$= E^{x}\{ e(QT(s))\}\, E^{x}\{ u(X(T(s))) \} \leq 2E^{x}\{ u(X(T(s))) \}$$

by independence and the second inequality in (4). Integrating as before we obtain

$$u(x) \le \frac{2}{V(r)} \int_{B(x,r)} u(y)\, dy \quad . \tag{11}$$

Since $B(x,r) \subset B(x_0,2r)$ and $u \ge 0$, (9) and (10) together yield

$$u(x) \le 2^{d+2}\, u(x_0) \ . \tag{12}$$

In particular we have proved that $u(x) < \infty$ if $\rho(x,x_0) < r$ and consequently we may interchange the roles of x_0 and x in the above. Since the number r is fixed independently of x, and K is compact, a familiar "chain argument" establishes the theorem. Indeed if N is the number of overlapping balls of fixed radius r which are needed to lead in a chain from any point to any other point in K, then the constant A in (3) may be taken to be $2^{(d+2)N}$. □

Corollary. If K is fixed and D is enlarged, the inequalities in (3) remain valid with the same constant A.

This is clear from the proof, and will be needed for the application in Theorem 3.1.

The following lemma plays a key role below. Its essential feature is that only the (Lebesgue) measure $m(E_n)$ of E_n, and not its shape or smoothness, is involved.

Lemma A. Let $\{E_n\}$ be sets with $m(E_n)$ decreasing to zero. Then we have for each $t > 0$:

(13) $$\lim_{n\to\infty} \sup_{x\in\bar{E}_n} P^x\{ \tau(E_n) > t \} = 0 .$$

For any constant Q we have

(14) $$\lim_{n\to\infty} \sup_{x\in\bar{E}_n} E^x\{ \exp(Q\tau(E_n)) \} = 1 .$$

Proof. We have for any E and $t > 0$:

(15) $$\sup_{x\in\bar{E}} P^x\{ \tau(E) > t \} \le \sup_{x\in\bar{E}} P^x\{ X(t) \in E \} \le \frac{m(E)}{(2\pi t)^{d/2}}$$

because the probability density of $X(t)$ is bounded by $(2\pi t)^{-d/2}$. This implies (13). Next we obtain from (15) followed by a Markovian iterative argument:

$$\sup_{x\in\bar{E}} P^x\{ \tau(E) > nt \} \le \left(\frac{m(E)}{(2\pi t)^{d/2}}\right)^n .$$

Therefore we have

$$E^x\{ \exp(Q(\tau(E))) \} \le \sum_{n=0}^{\infty} e^{Q(n+1)t} P^x\{ \tau(E) > nt \}$$

$$\le e^{Qt} \sum_{n=0}^{\infty} [e^{Qt} m(E) (2\pi t)^{-d/2}]^n .$$

Given Q, chose t so small that Qt is near zero. For this t, if $m(E)$ is small enough the infinite series above has a sum near 1. This proves (14). □

It follows from Theorem 1.1 that if $u \not\equiv \infty$ in D then $u < \infty$ in D. When $m(D) < \infty$, this result has a sharpening which is not valid in the usual analytical setting of Harnack inequalities, in which only

local boundedness can be claimed. The situation will be clarified in later sections when we relate the function u to a positive solution of the Schrödinger equation.

Theorem 1.2. Let D be a domain with $m(D) < \infty$, and let q and f be as in Theorem 1.1, but f be bounded as well as nonnegative. If $u(q,f;\cdot) \not\equiv \infty$ in D, then it is bounded in $\bar{D}$.

Proof. Let us remark that if $m(D) < \infty$, then $P^x\{ \tau_D < \infty \} = 1$ for all $x \in R^d$, so that we may omit "$\tau_D < \infty$" in the definition (2). Write u as before and let $\|q\| = Q$. Let K be a compact subset of D such that $m(E) < \delta$ where $E = D-K$, and where δ is so small that

$$\sup_{x\in E} E^x\{ \exp(Q\tau(E)) \} \le 1+\varepsilon \ . \tag{16}$$

This is possible by Lemma A. Note that E is open and $\tau_E \le \tau_D$. For $x \in \bar{E}$ let us put

$$\begin{aligned} u_1(x) &= E^x\{ e(\tau_D)\ f(X(\tau_D));\ \tau_E < \tau_D \}, \\ u_2(x) &= E^x\{ e(\tau_D)\ f(X(\tau_D));\ \tau_E = \tau_D \}. \end{aligned} \tag{17}$$

We have by the strong Markov property:

$$\begin{aligned} u_1(x) &= E^x\{ \tau_E < \tau_D;\ e(\tau_E)\ E^{X(\tau_E)}[e(\tau_D)\ f(X(\tau_D))] \} \\ &= E^x\{ \tau_E < \tau_D;\ e(\tau_E)\ u(X(\tau_E)) \}. \end{aligned} \tag{18}$$

On the set $\{ \tau_E < \tau_D \}$, we have $X(\tau_E) \in K$, and u is bounded on K by Theorem 1.1. Hence we have by (16) and (18):

(19) $$u_1(x) \le E^x\{ \exp(Q\tau_E) \} \|u\|_K \le (1+\varepsilon) \|u\|_K .$$

On the other hand, we have for $x \in \bar{E}$:

(20) $$u_2(x) \le E^x\{ e(\tau_E) f(X(\tau_E)) \}$$
$$\le E^x\{ e(\tau_E) \} \|f\| \le (1+\varepsilon) \|f\| .$$

Combining the last two inequalities we have

(21) $$u(x) \le (1+\varepsilon)(\|u\|_K + \|f\|).$$

Since $\bar{D}-\bar{E} \subset K$, (21) holds trivially for $x \in \bar{D}-\bar{E}$. Thus (21) holds for all $x \in \bar{D}$. □

It is clear how we can make more precise the dependence of ε in (21) on K, thereby giving an estimate of the global bound $\|u\|_{\bar{D}}$ in terms of a local bound $\|u\|_K$ and $\|f\|$. Theorem 1.2 is true without any condition on the smoothness of ∂D. In the probabilistic treatment of the Dirichlet problem a point z is said to be a regular boundary point iff $z \in \partial D$ and $P^z\{ \tau_D = 0 \} = 1$, namely iff z is regular for D^c. The equivalence of this definition of regularity with the classical definition based on the solvability of the boundary value problem is well known. The next result is an extension of the probabilistic solution to the Dirichlet problem (D,f) to the present setting when the Feynman-Kac functional e_q is attached to the Brownian motion process. It will be seen in §2 that this extension is tantamount to replacing the Laplacian operator Δ by the Schrödinger operator $\Delta+2q$. When $q \equiv 0$ the theorem below reduces to Dirichlet's first boundary value problem.

Theorem 1.3. Let D and q be as in Theorem 1.2, but $f \in b\mathcal{B}(\partial D)$. If z is a regular point of ∂D and f is continuous at z, then we have

$$\lim_{x \to z} u(x) = f(z). \tag{22}$$

Remark. Since u is defined in $\bar{D}$ it is natural that the variable x in (22) should vary in $\bar{D}$, and not just in D. This minor but nontrivial point is sometimes overlooked.

Proof. Without loss of generality we may suppose $f \geq 0$. Given $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\sup_{x \in R^d} E^x\{ e^{2QT_r} \} \leq 1 + \varepsilon \qquad \text{for } r \leq \delta; \tag{23}$$

$$\sup_{y \in B(z,2\delta) \cap (\partial D)} |f(y) - f(z)| \leq \varepsilon. \tag{24}$$

Let $x \in B(z,\delta)$, and $0 < r < \delta$. Write τ for τ_D and put

$$u_1(x) = E^x\{ T_r < \tau;\ e(\tau)\ f(X(\tau)) \},$$

$$u_2(x) = E^x\{ \tau \leq T_r;\ e(\tau)\ f(X(\tau)) \}.$$

It is well known that for each $t > 0$, $P^x\{ \tau > t \}$ is upper semicontinuous in $x \in R^d$. Since $P^z\{ \tau > t \} = 0$, it follows easily that

$$\lim_{x \to z} P^x\{ \tau > T_r \} = 0 \tag{25}$$

where $x \in \bar{D}$. We have by the strong Markov property:

$$u_1(x) = E^x\{ T_r < \tau;\ e(T_r)u(X(T_r)) \}.$$

Hence by Theorem 1.2 followed by Schwarz's inequality:

$$u_1(x) \le E^x\{ T_r < \tau;\ e^{QT_r} \} \|u\|_D$$

$$\le P^x\{ T_r < \tau \}^{\frac{1}{2}} E^x(e^{2QT_r})^{\frac{1}{2}} \|u\|_D .$$

Therefore $\lim_{x\to z} u_1(x) = 0$ by (23) and (25). Next we have by (24), since $X(\tau) \in B(z,2\delta)$ on $\{ \tau \le T_r \}$ under P^x:

$$|u_2(x)-E^x\{ \tau \le T_r;\ e(\tau)f(z) \}| \le E^x\{ \tau \le T_r;\ e^{QT_r} \}\varepsilon \le (1+\varepsilon)\varepsilon ;$$

and by (23):

$$|1-E^x\{ \tau \le T_r;\ e(\tau) \}| \le P^x\{ T_r < \tau \} + E^x\{ \tau \le T_r;\ e(\tau)-1 \}$$

$$\le P^x\{ T_r < \tau \} + E^x\{ e^{QT_r} \} - 1$$

$$\le P^x\{ T_r < \tau \} + \varepsilon.$$

Since ε is arbitrary, it follows from the above inequalities and (25) that $\lim_{x\to z} u_2(x) = f(z)$. Thus (22) is true. □

The intuitive content of Theorems 1.2 and 1.3 is this: the motion of the Brownian path in a domain is such that large positive values cancel large negative values of $q(X(t))$ so neatly that no after-effect is felt as it approaches the boundary, provided that cancellation is possible in an average sense, measured exponentially. Moreover the latter possibility is irrespective of the starting point of the path.

2. Schrödinger equation

Let D be a domain in R^d. We introduce the notation

$$(1) \qquad Q_t f(x) = E^x\{ t < \tau_D;\ f(X_t) \}$$

for $f \in b\mathcal{B}$. Then $\{ Q_t,\ t \geq 0 \}$ is the transition semigroup of the Brownian motion killed upon the exit from D. Let

$$(2) \qquad G_D f(x) = E^x\{ \int_0^{\tau_D} f(X_t)\, dt \}$$

where the right member is defined first for $f \geq 0$, then through $f = f^+ - f^-$ in the usual way, provided either $G_D f^+$ or $G_D f^-$ is finite. We shall be concerned only with the case where $G_D|f| < \infty$. Let $C^{(0)}(D)$ and $C^{(k)}(D)$, $k \geq 1$, denote respectively the classes of continuous and k times continuously differentiable functions on D. We write $f \in H(D)$ and say that f is Hölder continuous in D, iff for any compact subset C of D there exist two constants $\alpha > 0$ and M such that $|f(x)-f(y)| \leq M|x-y|^\alpha$ for x and y in C. For a proof of the following lemma see e.g. [6; Chapter 4, §§5-6].

Lemma B. If f is locally bounded in D and $G_D|f| < \infty$ then $G_D f \in C^{(1)}(D)$. If in addition $f \in H(D)$ then $G_D f \in C^{(2)}(D)$, and

$$(3) \qquad \Delta(G_D f) = -2f \ .$$

On the other hand if $f \in C^{(2)}(D)$ then

$$G_D(\Delta f) = -2f + h,$$

where h is harmonic in D.

Let $q \in b\mathcal{B}$ as in §1. The Feynman-Kac semigroup $\{ K_t, t \geq 0 \}$ is defined as follows:

$$(4) \qquad K_t f(x) = E^x\{ e_q(t) f(X_t) \}$$

for $f \in b\mathcal{B}$. Actually Feynman considered a purely imaginary q and Kac a nonpositive q. For our q the semigroup need not be submarkovian. It is known that its infinitesimal generator is $\frac{\Delta}{2} + q$ (see [3]). When $q \equiv 0$, of course $\{K_t\}$ reduces to the Brownian semigroup $\{P_t\}$. In this case the function u in (2) of §1 is harmonic in D, namely it satisfies the Laplace equation $\Delta u = 0$ there. Theorem 1.1 becomes a classical Harnack theorem for harmonic functions and Theorem 1.3 becomes Dirichlet's first boundary value problem. We are now going to show that for a general bounded q the function u satisfies the Schrödinger equation (5) below.

Theorem 2.1. Let D, q and f be as in the definition (2) of §1 except that f need not be nonnegative. If $u(q, |f|; \cdot) \not\equiv \infty$ in D, then $u(q,f;\cdot) \in C^{(1)}(D)$. If in addition $q \in H(D)$, then $u(q,f;\cdot)$ satisfies the equation

$$(5) \qquad (\Delta+2q)u = 0 \qquad \text{in } D \,.$$

Proof. Since the conclusions are local properties let us begin by localization. Writing u as before we see that it is locally bounded by Theorem 1.1. Let B be a small ball such that $\bar{B} \subset D$, and

$$(6) \qquad \sup_{x \in \bar{B}} E^x\{ \exp(Q\tau_B) \} < \infty \,.$$

We have for $x \in B$:

$$u(x) = E^x\{ e(\tau_B)\, u(X(\tau_B)) \} . \tag{7}$$

Comparing this with the definition of u we see that we have replaced (D,f) by (B,u), where B is bounded and satisfies (6), and u is bounded in $\bar{B}$. We need only prove the conclusions of the theorem for x in B. Reverting to the original notation we may therefore suppose that the domain D has the properties of B above, in particular $\tau_D < \infty$ and $E^x\{ e(\tau_D) \}$ is bounded in $\bar{D}$; and f is bounded. These conditions will be needed in the use of Fubini's theorem in the calculations which follow. [Warning: one must check the finiteness of the quantities below when q and f are replaced by $|q|$ and $|f|$; the former replacement is *not* trivial.] We write τ for τ_D and put for $0 \le s < t$:

$$e(s,t) = \exp(\int_s^t q(X(r))\, dr),$$

thus $e(\tau) = e(0,\tau)$. We have by the Markov property:

$$E^x\{ 1_{\{s<\tau\}}\, e(s,\tau)\, f(X_\tau) | F_s \} = 1_{\{s<\tau\}}\, u(X_s). \tag{8}$$

This relation is used in the first and last equations below:

$$\begin{aligned} (9)\quad & E^x\{ \int_0^t 1_{\{s<\tau\}}\, q(X_s)\, u(X_s)\, ds \} \\ &= E^x\{ \int_0^{t\wedge\tau} q(X_s)\, e(s,\tau)\, f(X_\tau)\, ds \} \\ &= E^x\{ [e(\tau) - e(t \wedge \tau, \tau)]\, f(X_\tau) \} \end{aligned}$$

$$= E^x\{ t < \tau;\ [e(\tau)-e(t,\tau)]\ f(X_\tau) \} + E^x\{ t \ge \tau;\ [e(\tau)-1]\ f(X_\tau)\}$$

$$= E^x\{ e(\tau)\ f(X_\tau) \} - E^x\{ t < \tau;\ u(X_t) \} - E^x\{ t \ge \tau;\ f(X_\tau) \}.$$

Now put

$$v(x) = E^x\{ f(X(\tau_D)) \}$$

for $x \in D$. Then v is the probabilistic solution of the Dirichlet problem (D,f) reviewed above; hence $\Delta v = 0$ in D. The last member of (9) may be written as

$$u(x) - Q_t u(x) - v(x) + Q_t v(x).$$

Since both u and v are bounded, and $\lim_{t\to\infty} Q_t 1 = 0$ because D is bounded, we have $\lim_{t\to\infty} Q_t(u-v) = 0$. We may therefore let $t \to \infty$ in the first member of (9) to obtain, with the notation of (2):

$$G_D(qu) = u-v \quad \text{in } D . \tag{10}$$

Since u as well as q is bounded, and $G_D 1 < \infty$ because D is bounded, we have $G_D(|qu|) < \infty$. Since v is harmonic it follows from (10) and Lemma B that $u \in C^{(1)}(D)$; if $q \in H(D)$ then $u \in C^{(2)}(D)$ and

$$\Delta u = \Delta v + \Delta G_D(qu) = -2qu$$

which is (5). □

Before going further let us recapitulate the essential part of Theorems 1.3 and 2.1, leaving aside the generalizations. Let D be a

bounded domain, $q \in b\mathcal{B}(D) \cap H(D)$, $f \in C^{(0)}(\partial D)$. Suppose that for some x_0 in D we have $u(q,1;x_0) < \infty$, then writing $u(x)$ for $u(q,f;x)$, we have $u \in C^{(2)}(D)$ and u is a solution of the Schrödinger equation $(\Delta+2q)u = 0$ in D. Furthermore $u(x)$ converges to $f(z)$ as x approaches each regular point z of ∂D. In particular if ∂D is regular then $u \in C^{(0)}(\bar{D})$. For $q \equiv 0$, u is the well known solution to the Dirichlet problem (D,f). Now in the latter case there is a converse as follows. Let $\varphi \in C^{(2)}(D) \cap C^{(0)}(\bar{D})$ and $\Delta\varphi = 0$ in D, then we have for all x in D:

$$\varphi(x) = E^x\{ \varphi(X(\tau_D)) \} .$$

This provides an extension of Gauss's average theorem for harmonic functions and implies the uniqueness of the solution to the Dirichlet problem. We proceed to establish corresponding results in the present setting.

The following lemma is stated for the sake of explicitness.

Lemma C. Let D, q, f be as in the definition of (2) of §1, except that f need not be nonnegative. If $u(q,|f|;\cdot) \not\equiv \infty$ in D, then we have for all $x \in D$ and $t \geq 0$:

$$E^x\{ e(\tau_D)\, f(X(\tau_D)) | F_t \} = e(t \wedge \tau_D)\, u(X(t \wedge \tau_D)) \quad . \tag{11}$$

Proof. We have

$$E^x\{1_{\{t<\tau_D\}} e(\tau_D) f(X(\tau_D)) | F_t\} = 1_{\{t<\tau_D\}} e(t)\, E^x\{e(t,\tau_D) f(X(\tau_D)) | F_t\}$$

$$= 1_{\{t<\tau_D\}}\, e(t)\, u(X_t)$$

by (8) (with s replaced by t). On the other hand,

$$E^x\{ 1_{\{t \ge \tau_D\}} e(\tau_D)\ f(X(\tau_D)) | \mathcal{F}_t \} = 1_{\{t \ge \tau_D\}} e(\tau_D)\ f(X(\tau_D))$$

because the trace of $\mathcal{F}_t$ on $\{ t \ge \tau_D \}$ contains the trace of $\mathcal{F}_{\tau_D}$ on $\{ t \ge \tau_D \}$. Now by Kellogg's theorem (see e.g. [6]) irregular points of ∂D form a polar set, hence $X(\tau_D)$ is a regular point of ∂D almost surely under P^x, $x \in D$; and consequently $u(X(\tau_D)) = f(X(\tau_D))$ by (2) of §1. Using this in the second relation above and adding it to the first relation we obtain (11). □

Theorem 2.2. Let D be an arbitrary domain and $q \in bC^{(0)}(D)$. Suppose that the function φ has the following properties:

$$\varphi \in C^{(2)}(D); \quad \varphi > 0 \text{ and } (\Delta+2q)\varphi = 0 \text{ in } D. \tag{12}$$

Then for any bounded subdomain E such that $\bar{E} \subset D$, we have

$$\forall\ x \in D:\ \varphi(x) = E^x\{ e(\tau_E)\ \varphi(X(\tau_E)) \}\ . \tag{13}$$

Proof. Although we can prove this result without stochastic integration, it is expedient to use Ito's formula. We have then in the customary notation:

$$d(e(t)\ \varphi(X_t)) = e(t)\{ \nabla\varphi(X_t)\ dX_t + (2^{-1}\Delta\varphi + q\varphi)(X_t)\ dt \} \tag{14}$$

where ∇ denotes the gradient and dX_t the stochastic differential (see e.g., [8]). The second term on the right side of (14) vanishes for $t < \tau_D$ by (12), and the first term is a local martingale. Since

$\bar{E}$ is compact and $\varphi \in C^{(2)}$ in a neighborhood of $\bar{E}$, we have

$$(15)\qquad \sup_{0 \le s \le t\wedge\tau_E} \| e(s)\, \nabla\varphi(X_s)\| \le e^{\|q\|t} \, \|\nabla\varphi\|_{\bar{E}} < \infty.$$

Hence (14) has the formal expression:

$$(16)\qquad e(t\wedge\tau_E)\,\varphi(X(t\wedge\tau_E)) - e(0)\varphi(X(0)) = \int_0^{t\wedge\tau_E} e(s)\, \nabla\varphi(X_s)\, dX_s = M_t,$$

say, where $\{ M_t, F_t, t \ge 0 \}$ is a martingale, under P^x for each $x \in D$, with $E^x(M_t) = 0$ for $t \ge 0$. Taking E^x in (16) we obtain

$$(17)\qquad \varphi(x) = E^x\{ e(t\wedge\tau_E)\,\varphi(X(t\wedge\tau_E)) \} \quad .$$

By enlarging E if necessary, we may assume that all boundary points of E are regular. By Theorem 1.3, the function v defined by

$$(18)\qquad v(x) = E^x\{ e(\tau_E)\,\varphi(X(\tau_E)) \}$$

is continuous on $\bar{E}$ and equals φ on ∂E. We can apply Lemma C with D and f replaced by E and φ to obtain

$$(19)\qquad E^x\{ e(\tau_E)\,\varphi(X(\tau_E)) \mid F_t \} = e(t\wedge\tau_E)\, v(X(t\wedge\tau_E)).$$

Since $\varphi > 0$, v cannot vanish; and since v is continuous, it is bounded below on $\bar{E}$. We conclude from (19) that

$$(20)\qquad e(t\wedge\tau_E) \le (\inf_{\bar{E}} v)^{-1}\, E^x\{ e(\tau_E)\,\varphi(X(\tau_E)) \mid F_t \} .$$

A simple consequence of (2) is that the family $e(t\wedge\tau_E)$, $t \ge 0$, of

random variables is uniformly integrable. We can thus let t tend to infinity in (17) to obtain

$$(21) \qquad \varphi(x) = E^x\{ e(\tau_E)\, \varphi(X(\tau_E)) \},$$

which shows (13) for $x \in \bar{E}$. Note that (13) is trivial for $x \in D-\bar{E}$.

Let us introduce the notation, for any $E \subset D$:

$$(22) \qquad u_E(x) = E^x\{ e(\tau_E) \} ;$$

in particular u_D is $u(q,1;\cdot)$ in our previous notation. Parts of Theorem 2.2 are important enough to be stated as corollaries.

Corollary 1. For each bounded domain (or open set) E such that $\bar{E} \subset D$, we have u_E bounded in D.

Corollary 2. If there exists a function φ satisfying the conditions in (12) and is furthermore bounded above and bounded away from zero, then u_D is bounded in $\bar{D}$. In particular, this is the case if D is bounded and $\varphi \in C^{(0)}(\bar{D})$.

Proof. For then we have from (13), for all $x \in D$:

$$(\inf_{\bar{D}} \varphi)\, u_E(x) \le \varphi(x) \le \sup_{\bar{D}} \varphi$$

and consequently u_E is uniformly bounded with respect to E. There exists a sequence of such subdomains E_n increasing to D so that

$\tau(E_n)$ increases to $\tau(D)$. Hence it follows by Fatou's lemma that

$$E^x\{ e(\tau_D) \} \le \varliminf_n E^x\{ e(\tau(E_n)) \}$$

is also bounded in D. Now it is easily shown that u_D is indeed bounded in $\bar{D}$. □

When $q \ge 0$, $\bar{D}$ is compact and ∂D is regular, the second assertion in Corollary 2 was proved by Khas'minskii [4] for strong Markov processes with strong Feller property and continuous paths. His method uses the Taylor series for $e(\tau_D)$, an iterative argument à la Picard, and a maximum principle. These depend essentially on the non-negativeness of q. The methods used in this paper can also be partially generalized to the class of processes considered by him, but the more difficult theorems such as Theorem 1.2 elude us at the moment.

Without the further assumption in Corollary 2 we cannot conclude that $u_D \not\equiv \infty$ in D; nor that (13) holds when E is replaced by D. The simplest example is in R^1 when $D = (0,\pi)$, $q = 1/2$, $\varphi(x) = \sin x$. The fact that $u_D \equiv \infty$ in this example will follow by contraposition from the results below. We need an easy but important lemma.

Lemma D. Let D be a domain with $m(D) < \infty$, and $u_D \not\equiv \infty$ in D. There exists a constant $c_0 > 0$ such that for all $x \in \bar{D}$:

$$u_D(x) \ge c_0(1 \vee u_E(x)) \tag{23}$$

where E is any set with $\bar{E} \subset D$. For all D and q such that $m(D) \le M < \infty$ and $\|q\| \le Q < \infty$, the constant c_0 depends only on M and Q.

Proof. It follows from (15) of §1 that for some t_0 depending only on M we have $P^x\{\tau_D \le t_0\} \ge 1/2$ for all $x \in \bar{D}$. We have then

$$u_D(x) \ge E^x\{ e(\tau_D);\ \tau_D \le t_0 \} \ge E^x\{ e^{-Qt_0};\ \tau_D \le t_0 \} \ge c_0,$$

where $c_0 = e^{-Qt_0}/2$. If $\bar{E} \subset D$ we have $\tau_E < \tau_D$ and so by the strong Markov property

$$u_D(x) = E^x\{ e(\tau_E)\ u_D(X(\tau_E)) \} \ge E^x\{ e(\tau_E) \}\ c_0 = u_E(x)\ c_0. \qquad \square$$

A simple consequence of Lemma D is the following converse to Corollary 2 above. If $m(D) < \infty$, and $u_D = \varphi$, then this φ satisfies (12); it is bounded above by Theorem 1.2 and bounded away from zero by Lemma D. In particular if D is bounded and ∂D is regular then $u_D \in C^{(0)}(\bar{D})$ by Theorem 1.3. The importance of this particular solution of the Schrödinger equation will become apparent in what follows.

Theorem 2.3. Let D be a domain with $m(D) < \infty$ and $u_D \not\equiv \infty$. Suppose that the function φ has the following properties:

$$(24) \qquad \varphi \in C^{(2)}(D) \cap bC^{(0)}(\bar{D}); \qquad (\Delta+2q)\varphi = 0 \quad \text{in } D.$$

Then we have

$$(25) \qquad \forall\ x \in D: \qquad \varphi(x) = E^x\{ e(\tau_D)\,\varphi(X(\tau_D)) \}.$$

Proof. Although φ is no longer positive, the proof of Theorem 2.2 needs no change up to (17) there. Now by (23), $u_E \le c_0^{-1}\ u_D < \infty$ in D; hence $e(\tau_E)$ is integrable under P^x, $x \in \bar{E}$. We may therefore

apply Lemma C to E with $f \equiv 1$ to obtain

$$(26)\quad E^x\{ e(\tau_E)|\mathcal{F}_t \} = e(t \wedge \tau_E)\, u_E(X(t \wedge \tau_E)) \geq e(t \wedge \tau_E)\, c_0$$

since $\inf_{\bar{E}} u_E \geq c_0$ by Lemma D. It follows that the family of random variables $\{ e(t \wedge \tau_E), t \geq 0 \}$ is uniformly integrable. Since φ is bounded continuous in $\bar{D}$ we may let $t \to \infty$ under E^x in (17) to deduce

$$(27)\qquad \varphi(x) = E^x\{ e(\tau_E)\varphi(X(\tau_E)) \}, \qquad x \in D.$$

Next we have, since $u_D < \infty$:

$$(28)\qquad E^x\{ e(\tau_D)|\mathcal{F}(\tau_E) \} = e(\tau_E)\, u_D(X(\tau_E)) \geq e(\tau_E)\, c_0 \quad ,$$

by Lemma D. Hence the family $\{e(\tau_E)\}$ is uniformly integrable as E ranges over all sets with closures contained in D. Let E_n be open bounded, $\bar{E}_n \subset E_{n+1} \subset D$ and $\bigcup_n E_n = D$. We obtain (25) by putting $E = E_n$ in (27) and passing to the limit as before.

Let us remark that under the assumptions of Theorem 2.3 it is possible to apply Ito's formula directly to D instead of E as we did, provided that we can deduce the boundedness of $\nabla\varphi$ in (14) from that of φ (and hence of $\Delta\varphi$). This would require some kind of global estimate of Schauder's type which might require stronger smoothness conditions.

If $q \equiv 0$ and D is bounded, the preceding theorem reduces to the classical result of the representation of a harmonic function by its boundary values. We are now ready to solve a similar problem for

the Schrödinger equation.

Theorem 2.4. Let D be as in Theorem 2.3, ∂D be regular, and $q \in H(D)$. For any $f \in bC^0(\bar{D})$ there is a unique φ satisfying (24) such that $\varphi \equiv f$ on ∂D. Indeed this φ is given by

$$(29) \qquad \varphi(x) = E^x\{ e(\tau_D)\, f(X(\tau_D)) \}, \qquad x \in D.$$

Proof. That this φ satisfies (24) is proved by Theorem 2.1 and Theorem 1.3. The uniqueness follows from Theorem 2.3. □

The relationship between the preceding results and those obtainable by the usual methods of partial differential equations will be discussed in a separate publication.

3. Further results

As a by-product of the methods used above which can be stated without mentioning probability, we give the following theorem about positive solutions of the Schrödinger equation. This will be needed in the next theorem.

Theorem 3.1. Let D be a domain and D_n be bounded domains such that $\bar{D}_n \subset D_{n+1}$ and $\bigcup_n D_n = D$. Let $q_n \in H(D_n)$ and $q_n \to q$ boundedly where $q \in H(D)$. Suppose that for each $n \geq 1$, there exists φ_n such that

$$(1) \qquad \varphi_n > 0, \qquad (\Delta + 2q_n)\varphi_n = 0 \quad \text{in } D_n.$$

Then there exists φ such that

(2) $\varphi > 0, \qquad (\Delta + 2q)\varphi = 0 \quad \text{in } D.$

Proof. Let Q be a common bound of all $\|q_n\|$ (and $\|q\|$). By Corollary 1 to Theorem 2.2, the existence of φ_n implies that for each $n \geq 2$ the function u_n defined by

(3) $$u_n(x) = E^x\{ e_{q_{n+1}}(\tau_{D_n}) \}$$

is bounded in $\bar{D}_n$. Choose any $x_0 \in D_1$ and put

(4) $$v_n(x) = u_n(x)/u_n(x_0) \quad .$$

According to the Corollary to Theorem 1.1, applied for $k > n$ to v_k on $\bar{D}_n$ and with $f \equiv 1$, there exists a constant $A_n > 0$ (depending only on D_{n+1}, D_n and Q) such that we have for all $k > n$:

(5) $$A_n^{-1} \leq \inf_{\bar{D}_n} v_k \leq \sup_{\bar{D}_n} v_k \leq A_n .$$

Define the measures below:

(6) $$\mu_k^{\pm}(dx) = q_{k+1}^{\pm}(x)v_k(x)dx, \quad \mu_k(dx) = \mu_k^{+}(dx) - \mu_k^{-}(dx) = q_{k+1}(x)v_k(x)dx$$

where $q_k = q_k^+ - q_k^-$ is the usual decomposition. The two sequences of measures $\{ \mu_k^{\pm}, k > n \}$ on D_n are vaguely bounded because

(7) $$\mu_k^{\pm}(D_n) \leq QA_n\, m(D_n), \qquad k > n.$$

We have by Theorem 2.1: $v_k \in C^{(2)}(D_k)$ and

(8) $$(\Delta+2q_{k+1})v_k = 0, \qquad \text{in } D_k .$$

Hence it follows from Lemma B that for $k > n$:

(9) $$v_k = G_{D_n}(q_{k+1}\, v_k) + h_k$$

where h_k is harmonic in D_n. We have by (9) and (5):

(10) $$\| h_n \|_{D_n} \le A_n + QA_n\, G_{D_n} 1 < \infty .$$

Hence by Harnack's theorem on harmonic functions, followed by a diagonal argument, there exists a sequence $\{k_j\}$ such that $\{h_{k_j}\}$ converges uniformly on each D_n, $n \ge 1$; and the limit is a harmonic function h in D. By (7), the sequence $\{k_j\}$ may be chosen so that both $\mu^+_{k_j}$ and $\mu^-_{k_j}$ converge vaguely on each D_n, $n \ge 1$. The bounds in (5) then imply that for every $f \in L^1(D)$ (not only for $f \in bC(D_n)$), $\int_{D_n} f d\mu_{k_j}$ converges as $j \to \infty$. Since D_n is bounded, we know that $G_{D_n}(x,dy) = g_{D_n}(x,y)\,dy$ where $g_{D_n}(x,\cdot) \in L^1(D_n)$. Here the function g_{D_n} is the Green's function for D_n. Therefore if we substitute k_j for k in (9), the first term on the right side converges as $j \to \infty$. Thus $\lim_j v_{k_j} = v$ exists on D_n for each $n \ge 1$, and we obtain from (9):

(11) $$v = G_{D_n}(qv) + h \qquad \text{in } D_n .$$

Since $v \ge A_n^{-1}$ in D_n by (5), $v > 0$ in D. Using Lemma C, we see first that $v \in C^{(1)}(D)$ since qv is bounded by (5), and then $v \in C^{(2)}(D)$ since $qv \in H(D)$; finally by (11):

(12) $$\Delta v = \Delta G_{D_n}(qv) + \Delta h = -2qv .$$

This is true in D_n for $n \geq 1$, hence also in D. We have proved the existence of φ in (2) since v is such a function. □

Several experts in partial differential equations were consulted about Theorem 3.1. Hans Weinberger said it could be proved by classical eigenvalue methods of solving the Dirichlet problem for strongly elliptic equations. S.T. Yau said it could be proved (and the bounded convergence of q_n generalized) by variational methods. N. Trudinger said he would prove it by using Harnack's inequalities as we did above. A related result is given in [5].

From our point of view, the interest of Theorem 3.1 lies in the observation that its probabilistic analogue is false. Namely, if D_n increases to D, the finiteness of u_{D_n} for all n does not imply the finiteness of u_D. Yet u_{D_n} satisfies (1) (with $q_n \equiv q$) for each n, and if $u < \infty$ it will satisfy (2). A counterexample is furnished by the example in R^1 given above.

Let

$$L_t f(x) = E^x\{\, t < \tau_D;\ e(t)f(X_t) \,\}$$

in analogy with (4) of §2. In the next theorem we relate the finiteness of u_D to several conditions on the semigroup $\{L_t\}$ just defined. The results hold for the u in (2) of §1 with bounded f, but we put $f \equiv 1$ for simplicity. Since D and q are fixed below we will write u for u_D, τ for τ_D and $e(t)$ for $e_q(t)$. Consider then the following statements:

(a) $u \not\equiv \infty$ in D;

(b) for every $x \in D$, we have

(13) $$\int_0^\infty L_t 1(x)dt < \infty ;$$

(c) there is at least one x_0 in D having the property that for every $\delta > C$, there exist $A(\delta)$ and $N(\delta)$ such that

(14) $$L_t\, 1(x_0) \le A(\delta)\, e^{\delta t} \qquad \text{for} \quad t \ge N(\delta) \quad ;$$

(d) there exists φ satisfying (12) of §2 above;

(e) for any bounded open set E such that $\bar{E} \subset D$, u_E is bounded in $\bar{E}$.

Theorem 3.2. For any domain D, (a) and (b) are equivalent; and (b) implies (c). If $q \in bC^{(0)}(D)$, then (d) implies (e). If $q \in H(D)$, then (c) implies (d).

Proof. By Theorem 1.1, (a) implies that $u(x) < \infty$ for all $x \in D$. For any $s > 0$ and $n \ge 1$ we have by the Markov property:

(15) $$E^x\{e(\tau);\ ns < \tau \le (n+1)s\} = E^x\{e(ns);\ ns < \tau;\ E^{X(ns)}[e(\tau);\ 0<\tau\le s]\}.$$

Using (15) of §1, we can choose s so that $C = \inf_{x\in R^d} P^x\{\tau \le s\} > 0$. Then it is trivial that for every $y \in D$, we have

(16) $$Ce^{-Qs} \le E^y\{\ e(\tau);\ 0 < \tau \le s\ \} \le e^{Qs}$$

where $Q = \|q\|$. It follows from (15) and (16) that (a) is equivalent to:

(17) $$\sum_{n=0}^{\infty} E^x\{\ e(ns);\ ns < \tau\ \} < \infty.$$

For $ns < t < (n+1)s$, we have $e^{-Qs}e(ns) \le e(t) \le e^{Qs}e(ns)$. Hence (17) is equivalent to (13) by an easy comparison. This proves that (a) and (b) are equivalent. Of course, (c) is a much weakened form of (b) via (17).

Next, for a fixed $\varepsilon > 0$, we have by (15) and the second inequality in (16) applied to the functional $e_{q-\varepsilon}$ instead of e_q:

$$E^x\{ e_{q-\varepsilon}(\tau);\ n < \tau \le n+1 \} \le e^{Q+\varepsilon}\ E^x\{ e_{q-\varepsilon}(n);\ n < \tau \}$$

$$= e^{Q+\varepsilon-n\varepsilon}\ E^x\{ e_q(n);\ n < \tau \} \quad .$$

If (c) is true, the last member above is bounded by $e^{Q+\varepsilon+n(\delta-\varepsilon)}\ A(\delta)$ for $n \ge N(\delta)$. We may choose $0 < \delta < \varepsilon$; then

$$(18) \qquad E^x\{e_{q-\varepsilon}(\tau)\} = \sum_{n=0}^{\infty} E^x\{e_{q-\varepsilon}(\tau);\ n<\tau\le n+1\} \le A(\delta)\ e^{Q+\varepsilon} \sum_{n=0}^{\infty} e^{n(\delta-\varepsilon)} < \infty .$$

Now assume $q \in H(D)$. We can then apply Theorem 2.1 to obtain $u_{q-\varepsilon} \in C^{(2)}(D)$ and

$$(19) \qquad (\Delta+2(q-\varepsilon))\ u_{q-\varepsilon} = 0.$$

Since this is true for every $\varepsilon > 0$, and $q-\varepsilon$ converges to q boundedly as $\varepsilon \to 0$, we can apply Theorem 3.1 to conclude that (d) is true. Finally, if $q \in bC^{(0)}(D)$, then (d) implies (e) by Corollary 1 to Theorem 2.2.

The conditions in (b) and (c) are meaningful in the spectral theory of the semigroup $\{L_t\}$, or its infinitesimal generator the Schrödinger operator. At least when (14) holds for all x in D, this interpretation should yield (d) in some sense as a result on the

point spectrum of the operator. It is not clear under what precise conditions the classical approach will confirm the results above obtained by probabilistic methods.

References

[1] K.L. CHUNG. On stopped Feynman-Kac functionals. *Séminaire de Probabilités XIV (Univ.Strasbourg).* Lecture Notes in Math. *784*, Springer-Verlag, Berlin, 1980.

[2] K.L. CHUNG and K.M. RAO. Sur la theorie du potentiel avec la fonctionelle de Feynman-Kac. *C.R. Acad. Sci. Paris 210* (1980).

[3] M. KAC. On some connections between probability theory and differential and integral equations. *Proc. Second Berkeley Symposium on Math. Stat. and Probability*, 189-215. University of California Press, Berkeley, 1951.

[4] R.Z. KHAS'MINSKII. On positive solutions of the equation Au + Vu = 0. *Theo. Prob. Appl. 4* (1959), 309-318.

[5] W.F. MOSS and J. PIEPENBRINK. Positive solutions of elliptic equations. *Pac. J. Math. 75* (1978), 219-226.

[6] S. PORT and C. STONE. *Brownian Motion and Classical Potential Theory.* Academic Press, New York, 1978.

[7] B. SIMON. *Functional Integration and Quantum Physics.* Academic Press, New York, 1979.

[8] D.W. STROOCK and S.R.S. VARADHAN. *Multidimensional Diffusion Processes.* Springer-Verlag, Berlin, 1979.

[9] E.B. DYNKIN. *Markov Processes.* Springer-Verlag, Berlin, 1965.

K.L. Chung
Department of Mathematics
Stanford University
Stanford, CA 94305, U.S.A.

K.M. Rao
Department of Mathematics
University of Florida
Gainesville, FL 32611, U.S.A.

Math. Ann. 264, 73–79 (1983)

With kind permission of Springer Science and Business Media.

Mathematische Annalen

Extension of Domains with Finite Gauge

K. L. Chung*, R. Durrett**, and Zhongxin Zhao

Department of Mathematics, Stanford University, Stanford, CA 94305, USA

Let $\{X_t\}$ be the Brownian motion in R^d, $d \geqq 1$; E^x and P^x denote respectively the expectation and probability for the process with $X_0 = x$. Let D be a domain in R^d, $d \geqq 2$, with $m(D) < \infty$ where m is the Lebesgue measure in R^d. All given sets and functions below are Borel measurable. For a bounded function q in R^d, and positive ($\geqq 0$) f on ∂D, we define

$$u(D, q, f; x) = E^x\{e_q(\tau_D) f(X(\tau_D))\} \tag{1}$$

where

$$e_q(t) = \exp\left(\int_0^t q(X_s)ds\right)$$

and

$$\tau_D = \inf\{t > 0 | X(t) \notin D\}.$$

It is proved in [4] that if $u(D, q, f; \cdot) \not\equiv \infty$ in D, then it is bounded in $\bar{D}$. We call $u(D, q, 1; \cdot)$ the *gauge* for (D, q). Since q is fixed in this paper but D will vary, we will denote the gauge by u_D, and say it is finite when $u_D \not\equiv \infty$ in D. We write also $\|u_D\|$ for $\sup_{x \in \bar{D}} u_D(x)$. The importance of the gauge is evident from the results in [4]. In this paper we study the question: if u_D is finite, can we enlarge D to a domain G so that u_G is still finite? First, we prove that we can always add a finite number of balls centered on ∂D to get such a G. But as we add more and more such balls, their radii may have to shrink to zero so that it is not always possible to cover the entire boundary of D. In fact, we shall give a simple example of a regular domain D with $u_D < \infty$, such that if part of D is added, the resulting domain G may have $u_G = \infty$. However, if D satisfies a uniform cone condition, in particular, if D is Lipschitzian, then there exists a domain $G \supset \bar{D}$ such that $u_G < \infty$. A discussion of these results from the point of view of eigenvalues follows the theorems.

* Research supported in part by NSF grant MCS-80-01540 at Stanford University

** Alfred P. Sloan Fellow; research supported in part by NSF grant MCS 80-02732 at University of California, Los Angeles

Let $z \in \partial D$ and $B(z,\varepsilon)$ be the ball with center z and radius ε. We write B_ε for $B(z,\varepsilon)$ below, and put

$$\varphi_\varepsilon(x) = u(D, q, 1_{B_\varepsilon}; x). \tag{2}$$

If $u_D < \infty$, then $\lim_{\varepsilon \to 0} \varphi_\varepsilon(x) = 0$ for all $x \in D$, because the singleton $\{z\}$ is a polar set. That is why we have supposed $d \geqq 2$. Moreover, $\varphi_\varepsilon \in C^{(1)}(D)$. Hence the convergence of φ_ε as $\varepsilon \downarrow 0$ is uniform in each compact subset of D by Dini's theorem. This is not sufficient for our later application, and we need the strengthening given below.

Lemma. *Suppose $u_D < \infty$. Let A be a compact subset of $\bar{D}$ and $z \notin A$. Then $\lim_{\varepsilon \to 0} \varphi_\varepsilon(x) = 0$ uniformly for $x \in A$.*

Proof. There is $\varepsilon > 0$ such that A is disjoint from $\overline{B(z,\varepsilon)}$. Let $x_0 \in A$ and fix a number $r > 0$ so that $r < \varrho(A, \overline{B(z,\varepsilon)})$, where ϱ denotes the distance, and also such that

$$\sup_{x \in B(x_0, r)} E^x\{\exp(Q\tau_{B(x_0, r)})\} < \infty \tag{3}$$

where $Q = \sup_x |q(x)|$. Writing τ_r for $\tau_{B(x_0, r)}$, we have by the strong Markov property, for each $x \in B(x_0, r)$:

$$\varphi_\varepsilon(x) = E^x\{\tau_r < \tau_D; e_q(\tau_r)\varphi_\varepsilon(X(\tau_r))\}. \tag{4}$$

Put $\tilde{\varphi}_\varepsilon = 1_D \varphi_\varepsilon$ and define for $x \in B(x_0, r)$:

$$\psi_\varepsilon(x) = E^x\{e_q(\tau_r)\tilde{\varphi}_\varepsilon(X(\tau_r))\} = u(B(x_0, r), q, \tilde{\varphi}_\varepsilon; x).$$

Since $\varphi_\varepsilon \leqq u_D$, φ_ε is bounded in $\bar{D}$ and $\tilde{\varphi}_\varepsilon$ is bounded in R^d. Now (3) implies that $u_{B(x_0, r)} < \infty$, hence $\psi_\varepsilon \in C^{(1)}(B(x_0, r))$ by Theorem 2.1 of [4]. By Dini's theorem, $\lim_{\varepsilon \to 0} \psi_\varepsilon = 0$ uniformly in $B(x_0, r/2)$. Since $\varphi_\varepsilon \leqq \psi_\varepsilon$, we have $\lim_{\varepsilon \to 0} \varphi_\varepsilon = 0$ uniformly in $B(x_0, r/2)$. This being true for every $x_0 \in A$, and A being compact, the lemma follows.

Theorem 1. *Let D be a domain in R^d, $d \geqq 2$, with $m(D) < \infty$ and $u_D < \infty$. For any $z \in \partial D$ there exists $\varepsilon > 0$ such that if $G = D \cup B(z,\varepsilon)$, then $u_G < \infty$. Furthermore, for any $\delta > 0$ there exists $\varepsilon(z,\delta)$ such that $\|u_G\| < \|u_D\| + \delta$ if $\varepsilon < \varepsilon(z,\delta)$.*

Proof. Given $0 < \delta < 1$, let η be such that

$$\sup_{x \in B(z, \eta)} E^x\{\exp(Q\tau_{B(z,\eta)})\} < 1 + \delta. \tag{5}$$

Observe that $B(z,\eta) \cap D$ may not be connected (see the blackened area in Fig. 1). Let

$$A = (\partial B(z,\eta)) \cap D.$$

We apply the lemma to find ε so that $0 < \varepsilon < \eta$ and

$$\sup_{x \in A} \varphi_\varepsilon(x) < \delta. \tag{6}$$

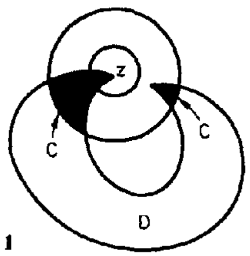


Fig. 1

Now put

$$C = G \cap B(z, \eta)$$
$$F = (\partial D) \cap B(z, \varepsilon)$$

where G is given in the statement of the theorem. Let $x \in D \cap B(z, \varepsilon)$. We shall prove that $u(G, q, 1; x) < \infty$. Then u_G will be finite by Theorem 1.2 of [4], as reviewed above.

The method of proof is similar to that of Theorem 1 in [1]*, which treats the case of a half line (instead of the D here) in R^1. It is somewhat more complicated owing to the geometry of R^d. Define $T_0 = 0$, and for $n \geqq 1$:

$$T_{2n-1} = T_{2n-2} + \tau_C \circ \theta_{T_{2n-2}},$$
$$T_{2n} = T_{2n-1} + \tau_D \circ \theta_{T_{2n-1}},$$
$$R_n = T_n \wedge \tau_G .$$

On $\{T_{2n-1} < \tau_G\}$, we have $X(T_{2n-1}) \in A$; on $\{T_{2n} < \tau_G\}$, we have $X(T_{2n}) \in F$. On $\{T_n < \tau_G\}$, $T_n < T_{n+1}$. Let $T_\infty = \lim_n \uparrow T_n$. On $\{T_\infty < \infty\}$, the path of the Brownian motion undergoes infinitely many oscillations of distance exceeding $(\eta - \varepsilon)/2$ before the time T_∞, since $\varrho(F, A) = \eta - \varepsilon$. The continuity of paths implies that $T_\infty = \infty$ a.s. (almost surely). Since $\tau_G < \infty$ a.s., it follows that there exists $n \geqq 1$ such that $T_{n-1} < \tau_G \leqq T_n$. But both sets C and D are subsets of G, and T_n is either an exit time from C or an exit time from D, the last inequalities entail that $\tau_G = T_n$, namely $\tau_G = R_n$. Hence if we define

$$N = \min\{n \geqq 0 | R_n = \tau_G\},$$

then $N < \infty$ a.s.

It follows from (5) that

$$\sup_{x \in F} E^x\{e(\tau_C)\} < 1 + \delta . \tag{7}$$

Applying the strong Markov property repeatedly to T_n, $n \geqq 1$, and using the estimates (6) and (7), we obtain

$$\begin{aligned} E^x\{e(\tau_G); N = 2n-1\} &\leqq [(1+\delta)\delta]^{n-1}(1+\delta); \\ E^x\{e(\tau_G); N = 2n\} &\leqq [(1+\delta)\delta]^{n-1}(1+\delta)\|u_D\| . \end{aligned} \tag{8}$$

* There is a minor error on p. 351. Replace the definition of S by $\tau_a \wedge \tau_c$, and put $N = \min\{n \geqq 0 | T_{2n+1} = \tau_c\}$

Let $(1+\delta)\delta<1$. Adding up (8) over $n\geqq 1$, we obtain

$$u_G(x)\leqq[1-(1+\delta)\delta]^{-1}(1+\delta)(1+\|u_D\|)<\infty.$$

In the above we have taken $x\in D\cap B(z,\varepsilon)$. A similar argument works for any $x\in D\backslash B(z,\varepsilon)$. Indeed, a slightly more refined argument shows that by taking ε small enough, we can make $\|u_G\|$ as near to $\|u_D\|$ as we wish. To see this let $B=B(z,\eta)$ and replace (7) by the following, for $x\in F$:

$$E^x\{e(\tau_B)1_A(X(\tau_B))\}<(1+\delta)\theta_x,$$
$$E^x\{e(\tau_B)1_{\partial B-A}(X(\tau_B))\}<(1+\delta)(1-\theta_x);$$

where θ_x is the ratio of the spherical area of A to the total area of ∂B. Although θ_x varies with x on F, by taking ε small enough in comparison with η, we can make $\theta'<\theta_x<\theta$ for all $x\in F$ and $\theta-\theta'<\delta$. The estimates on the right sides of (8) are then replaced by

$$[(1+\delta)\theta\delta]^{n-1}(1+\delta)(1-\theta') \quad\text{and}\quad [(1+\delta)\theta\delta]^{n-1}(1+\delta)\theta\|u_D\|,$$

respectively. The result is that

$$u_G(x)\leqq[1-(1+\delta)\theta\delta]^{-1}(1+\delta)(1-\theta'+\theta\|u_D\|), \tag{9}$$

for $x\in D\cap B(z,\delta)$; and similarly

$$u_G(x)\leqq[1-(1+\delta)\delta\theta]^{-1}[\|u\|_D+\delta(1+\delta)(1-\theta')], \tag{10}$$

for $x\in D\backslash B(z,\delta)$. Observe that $\|u_D\|\geqq 1$ because $u_D(z)=1$ if $z\in\partial D$ and z is regular (for D^c). It follows that as $\delta\downarrow 0$, the right member of (9) approaches $\|u_D\|$ as well as that of (10). Thus $\|u_G\|$ approaches $\|u_D\|$ as claimed.

We come next to the example mentioned in the introduction.

Example. Let $q\equiv 1$ in R^d. It is well known that there exists a number r_1 such that

$$E^0\{e^{\tau B(0,r)}\}<\infty$$

if and only if $r<r_1$. Let $r_3<r_2<r_1$, $B_i=B(o,r_i)$, and $C=B_1-\bar{B}_3$. We may make r_1-r_3 so small that

$$\sup_{x\in C}u(C;1,1;x)<2.$$

Since $u(B_2,1,1;x)$ is continuous in B_2 and $u(B_2,1,1_A;x)$ decreases to zero as $A\downarrow\emptyset$, where A is an arc on the circle ∂B_2, we may make A small enough that

$$\sup_{x\in B_3}u(B_2,1,1_A;x)<\tfrac{1}{3}.$$

Let L be a closed line segment connecting ∂B_1 and ∂B_2-A. Now define

$$D=B_1-[(\partial B_2-A)\cup L].$$

Thus D is a simply connected domain (L is used only to ensure this). It is easy to see that D is a regular domain. If we denote $(\partial B_2-A)\cup L$ by E, then $E\subset\partial D$ and $D\cup E=B_1$, and $u_{B_1}=\infty$ by the definition of r_1. Clearly, for any domain $G\supset\bar{D}$ we have $u_G=\infty$. It remains to show that $u_D<\infty$.

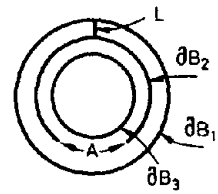


Fig. 2

The proof is similar to and simpler than that of Theorem 1.

Let $T_0=0$, and for $n\geqq 1$:

$$T_{2n-1}=T_{2n-2}+\tau_{B_2}\circ\theta_{T_{2n-2}},$$
$$T_{2n}=T_{2n-1}+\tau_C\circ\theta_{T_{2n-1}},$$
$$N=\min\{n\geqq 0\,|\,T_n=\tau_D\}.$$

Then $N<\infty$ a.s. We have

$$E^0\{e(\tau_D);N=2n-1\}=(2/3)^{n-1}u_{B_2},$$
$$E^0\{e(\tau_D);N=2n\}=(2/3)^n.$$

Hence $E^0\{e(\tau_D)\}<\infty$.

The cone condition is well known in Dirichlet's boundary value problem. Let us denote by $C(z,\theta)$ the cone with vertex z and relative angle θ; namely the intersection of the cone with the sphere $\partial B(z,1)$ has an area in the ratio $\theta:1$ to the total area of the sphere. A domain D is said to satisfy a cone condition at $z\in\partial D$ iff there exist $a>0,\theta>0$, so that $C(z,\theta)\cap B(z,a)\subset D^c$. If so, z is regular for D^c (see, e.g., [2] where a weaker cone condition is given). The condition is uniform iff the numbers a and θ can be taken to be the same for all $z\in\partial D$. If D satisfies a cone condition at every $z\in\partial D$ (not necessarily uniform), then it follows from Lebesgue's density theorem for measurable sets that $m(\partial D)=0$. It is easy to see that a bounded Lipschitz domain satisfies a uniform cone condition. We owe the last two remarks to N. Falkner.

Theorem 2. *Let D be a bounded domain with $u_D<\infty$, and satisfying a uniform cone condition. Then there exists a domain G containing $\bar{D}$ with $u_G<\infty$.*

Proof. Put

$$G=\{x\in R^d\,|\,\varrho(x,D)<\varepsilon\},\tag{11}$$

for some $\varepsilon>0$ to be determined later. Given $\delta>0$, let $0<\varepsilon_0<a$ such that

$$E^x\{\exp(Q\tau_{B(x,\varepsilon_0)})\}<1+\delta;\tag{12}$$

the number in (12) being independent of x. We decrease ε_0 if necessary so that for any domain E with $\bar{E}\subset D$ and $\varrho(\bar{E},\partial D)=\varepsilon_0$, and any G defined in (11) with $\varepsilon<\varepsilon_0$, $m(G-\bar{E})$ is small enough to satisfy the following conditions, where $C=G-\bar{E}$:

$$\sup_{x\in C}E^x\{\exp(Q\tau_C)\}<1+\delta;\tag{13}$$

$$\sup_{x\in\partial E}u_D(x)<1+\delta.\tag{14}$$

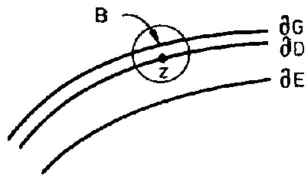


Fig. 3

Since $m(\partial D)=0$, $m(C)\to 0$ as $\varepsilon<\varepsilon_0\downarrow 0$; hence (13) is satisfied for small enough ε_0 by Lemma A of [4]. Since D is regular, u_D is continuous on $\bar{D}$ with boundary value one by Theorem 1.3 of [4]. Hence (14) is also satisfied for small enough ε_0.

Now let $\varepsilon_1<\varepsilon_0$. Then $\partial B(z,\varepsilon_1)\cap D^c$ has relative area $>\theta$ by the uniform cone condition, since $\varepsilon_1<$a. For any $0<\theta'<\theta$, by shrinking the angle of the cone, we obtain a subset $S(z,a)$ of $\partial B(z,a)\cap D^c$ which has relative area $>\theta'$, but with the additional property that

$$0<\varrho(S(z,a),D)<\varepsilon_0\,. \tag{15}$$

This number $\varrho(S(z,a),D)$ may be taken to be the same for all $z\in\partial D$, and we use it as the ε in the definition (11) of G. This choice of ε makes G disjoint from $S(z,\varepsilon_1)$, so that

$$\partial B(z,\varepsilon_1)\cap G^c \text{ has relative area } >\theta'\,. \tag{16}$$

At the same time, since $\varrho(\bar{E},\partial D)=\varepsilon_0>\varepsilon_1$, we have

$$B(z,\varepsilon_1)\cap\bar{E}=\emptyset\,. \tag{17}$$

The geometrical preparation is now complete, and we are ready for the key estimate below.

Fix $z\in\partial D$, and write $B=B(z,\varepsilon_1)$. We shall prove that $u_G(z)<\infty$. Under P^z, $\partial C=(\partial G)\cup(\partial E)$, and $\{\tau_C<\tau_G\}=\{X(\tau_C)\in\partial E\}\subset\{\tau_B<\tau_C\}\cap\{X(\tau_B)\in G\}$. Hence the first inequality below follows from the strong Markov property:

$$\begin{aligned}
E^z\{\exp(Q\tau_C);X(\tau_C)\in\partial E\} &\leqq E^z\{\tau_B<\tau_C;\exp(Q\tau_B)1_G(X(\tau_B))E^{X(\tau_B)}[\exp(Q\tau_C)]\}\\
&\leqq E^z\{\exp(Q\tau_B)1_G(X(\tau_B))\}\,(1+\delta)\\
&=E^z\{\exp(Q\tau_B)\}P^z\{X(\tau_B)\in G\}\,(1+\delta)\\
&\leqq(1+\delta)^2(1-\theta')\,.
\end{aligned}$$

The second inequality above follows from (13); the third from (12), (16), and spherical symmetry; the equality follows from the stochastic independence of τ_B and $X(\tau_B)$ under P^z. Since δ is arbitrarily small, the resulting bound may be made strictly less than one, which will suffice.

Define $T_0=0$, and for $n\geqq 1$:

$$\begin{aligned}
T_{2n-1}&=T_{2n-2}+\tau_C\circ\theta_{T_{2n-2}}\,,\\
T_{2n}&=T_{2n-1}+\tau_D\circ\theta_{T_{2n-1}}\,,\\
N&=\min\{n\geqq 1\,|\,T_{2n-1}=\tau_G\}\,.
\end{aligned}$$

Then $N < \infty$ a.s. We have

$$E^z\{e(\tau_G); N = 2n-1\} \leqq [(1+\delta)^3(1-\theta')]^{n-1}(1+\delta),$$

where the third $1+\delta$ factor comes from (14), when the path moves from ∂E back to ∂D. Choose δ so that $(1+\delta)^3(1-\theta') < 1$. It follows by summing over n that $u_G(z) < \infty$, indeed, $u_G(z)$ is arbitrarily near θ^{-1} for sufficiently small δ, since θ' may be arbitrarily near θ. For any $x \in G$, the same argument yields a bound for $u_G(x)$ arbitrarily near $\|u_D\|\theta^{-1}$. Thus, there exists G containing $\bar{D}$ such that $\|u_G\|$ is arbitrarily near $\|u_D\|\theta^{-1}$.

We do not know whether the last inequality can be improved as in the case of Theorem 1.

The results above about enlarging the domain while keeping the gauge finite are intimately connected with the variation of eigenvalues with the domain. Consider the following eigen equation:

$$\begin{gathered}\left(\frac{\Delta}{2} + q\right)\varphi = \lambda\varphi \text{ in } D;\\ \varphi = 0 \text{ on } \partial D.\end{gathered} \tag{18}$$

It is known that there exists a maximum eigenvalue $\lambda_1(D)$ for which (18) is solvable with $\varphi \in C^0(\bar{D}) \cap C^{(2)}(D)$, provided that q is Hölder continuous in D (as well as bounded). If D is regular, then it is shown in [3] and [6] by different methods that $u_D < \infty$ is equivalent to $\lambda_1(D) < 0$. Now there is a "principle" in classical analysis which asserts that $\lambda_1(D)$ varies continuously with D (at least when $q \equiv 0$). It should follow from this that if $\lambda_1(D) < 0$, then for a domain G "slightly larger" than D we should have $\lambda_1(G) < 0$. However, it is not clear under what precise conditions the said principle is valid. Conditions given in [5] are very strong in comparison with that used in Theorem 2 above, whereas Theorem 1 requires no condition on D except $m(D) < \infty$. These results are proved without any reference to eigenvalues.

References

1. Chung, K.L.: On stopped Feynman-Kac functionals, Séminaire de probabilités XIV (Univ. Strasbourg). Lecture Notes in Mathematics, No. 784. Berlin, Heidelberg, New York: Springer 1980
2. Chung, K.L.: Lectures from Markov processes to Brownian motion. Grundlehren der Mathematik, Nr. 249. Berlin, Heidelberg, New York: Springer 1982
3. Chung, K.L., Li, P.: Comparison of probability and eigenvalue methods for the Schrödinger equation. Adv. Math. (to appear)
4. Chung, K.L., Rao, K.M.: Feynman-Kac functional and the Schrödinger equation, Seminar on Stochastic Processes, pp. 1–29. Basel, Boston, Stuttgart: Birkhäuser 1981
5. Courant, R., Hilbert, D.: Methoden der mathematischen Physik, Vol. 1. Berlin: Springer 1931
6. Zhongxin Zhao: Local Feynman-Kac semigroup. J. Sys. Sci. Math. Sci. 2, 314–326 (1982)

Received 17 December 1982

Added in proof. Theorems 1 and 3 can be extended to the class of unbounded functions q, considered by Aizenman and Simon in Comm. Pure Appl. Math. 35, 209–271 (1982)

Reprinted with kind permission.

Ann. Inst. Henri Poincaré,

Vol. 20, n° 4, 1984, p. 349-351. *Probabilités et Statistiques*

The lifetime of conditional Brownian motion in the plane

by

Kai Lai CHUNG (*)

Department of Mathematics, Stanford University, CA 94305, USA

SUMMARY. — In this note I give a short and perspicacious proof of a recent remarkable result due to Cranston and McConnell [*3*].

RÉSUMÉ. — Cette note est consacrée à une démonstration courte d'un résultat remarquable et récent de Cranston et McConnell [*3*].

Let D be a bounded domain in R^d, $d \geq 1$; H(D) the class of strictly positive harmonic functions in D; $\mathrm{X} = \{ \mathrm{X}_t, t \geq 0 \}$ the standard Brownian motion in R^d; $\tau_\mathrm{B} = \inf \{ t > 0 : \mathrm{X}_t \notin \mathrm{B} \}$ for any Borel set B; m the Lebesgue measure in R^d; E_h^x the expectation associated with the h-conditioned Brownian motion starting at $x \in \mathrm{D}$.

THEOREM. — *Let* $d = 2$. *There exists a constant* C *depending only on* D *such that*

$$\sup_{\substack{x \in \mathrm{D} \\ h \in \mathrm{H}(\mathrm{D})}} \mathrm{E}_h^x \{ \tau_\mathrm{D} \} \leq \mathrm{C}m(\mathrm{D}) . \tag{1}$$

We begin by stating explicitly the case where $h \equiv 1$, namely unconditioned Brownian motion, for a general Borel set B in R^d, $d \geq 1$.

(*) Research supported in part by NSF grant MCS83-01072 at Stanford University.

LEMMA. — *We have*

$$\sup_{x \in D} E^x \{ \tau_B \} \leq A_d m(D)^{2/d}$$

where

$$A_d = \frac{1}{2\pi d^2}(d+1)^{\frac{2(d+1)}{d}}$$

This lemma can be proved by an elementary method using only the strong Markov property of X and the form of its transition density. It is generalizable and adaptable to similar estimates; see [*1*], p. 148 ff.

As the first simplification in the proof of the theorem, we deal directly with a general h in H(D). This spares us some unnecessary « hard theory », such as the famous Martin representation, and the behavior of a minimal harmonic function at the boundary. Cf. Lemma 2.2 in [*3*], which is actually a result due to Doob. Thus for any $h \in H(D)$, we put for clarity:

$$(2) \qquad Y(t) = \begin{cases} \left(\frac{1}{h}\right)(X_t) & \text{for} \quad 0 \leq t < \tau_D, \\ 0 & \text{for} \quad \tau \geq \tau_D. \end{cases}$$

It is a basic idea in h-conditioning that $\{ Y_t, \mathscr{F}_t, t \geq 0 \}$ is a super-martingale, where $\{ \mathscr{F}_t \}$ is the natural filtration of $\{ X_t \}$. Let $0 < a < b < \infty$; let $D'[a, b]$ and $U'[a, b]$ denote respectively the number of downcrossings and upcrossings of $[a, b]$ by $\{ Y_t, t \geq 0 \}$. Then we have for any $x \in D$:

$$(3) \qquad E^x \{ D'[a, b] \} \leq \frac{b}{b-a}; \qquad E^x \{ U'[a, b] \} \leq \frac{a}{b-a}.$$

For the first inequality (due to G. A. Hunt), see e. g. [*2*], p. 341; the second does not follow trivially from the first, but both follow from Dubins's inequalities (*loc. cit.*). Taking reciprocals, we deduce that if $D[a, b]$ and $U[a, b]$ denote the corresponding numbers for $\{ h(X_t), t \geq 0 \}$, then

$$(4) \qquad E^x \{ U[a, b] \} \leq \frac{a}{b-a}; \qquad E^x \{ D[a, b] \} \leq \frac{b}{b-a}.$$

We now define for any $x_0 \in D$:

$$C_n = \{ x \in D : h(x) = 2^n h(x_0) \},$$
$$D_n = \{ x \in D : 2^{n-1} h(x_0) < h(x) < 2^{n+1} h(x_0) \},$$

where n is an integer. Furthermore, we denote by N_n the total number of times a path moves from inside D_n to outside D_n. If it starts from C_n, this

can be done either by a downcrossing of $[2^n h(x_0), 2^{n-1}h(x_0)]$, or an upcrossing of $[2^n h(x_0), 2^{n+1}h(x_0)]$. Hence we have by (5):

$$\sup_{x\in C_n} E_h^x\{N_n\} \le \frac{2^{n-1}}{2^n - 2^{n-1}} + \frac{2^{n+1}}{2^{n+1} - 2^n} = 3. \tag{6}$$

Next, it is plain that for any $x \in C_n$:

$$E_h^x\{\tau_{D_n}\} = \frac{1}{h(x)} E^x\left\{\int_0^{\tau_{D_n}} h(X_t)dt\right\} \le 2E^x\{\tau_{D_n}\}. \tag{7}$$

It remains to add up all the crossings, *without ordering*. But we must not forget that a path may leave D before completing the last crossing. In this case 1 must be added to N_n in the counting. Therefore our final estimate is as follows:

$$E_h\{\tau_D\} \le (3+1)\sum_{n=-\infty}^{\infty} 2\sup_{x\in C_n} E^x\{\tau_{D_n}\} \le 8\,A_d \sum_{n=-\infty}^{\infty} m(D_n)^{\frac{2}{d}} \tag{8}$$

by (6), (7) and the lemma. For $d = 2$, this yields (1) with $C = 8A_d$.

REFERENCES

[1] K. L. CHUNG, *Lectures from Markov Processes to Brownian motion, Grundlehren der mathematischen Wissenschaft*, t. **249**, Springer-Verlag, 1982.

[2] K. L. CHUNG, *A course in Probability Theory*, 2nd edition, Academic Press, 1974.

[3] M. CRANSTON and T. R. MCCONNELL, The lifetime of conditioned Brownian motion, *Z. Wahrsch. Verw. Gebiete*, t. **65**, 1983, p. 1-11.

(Manuscrit reçu le 16 mars 1984)

Seminar on Stochastic Processes, 1983
Birkhäuser, Boston, 1984

DUALITY UNDER A NEW SETTING

by

K. L. Chung*, Ming Liao* and K. M. Rao

0. Introduction

This is a continuation of [2]. The developments there are complicated by an exceptional set denoted by Z (see [2], p. 179). It is shown that Z is a polar set under the conditions there if and only if Hunt's Hypothesis (B) holds (see [2], p, 192). In this paper a set of sufficcient conditions on the potential kernel will be given for the absence of Z. These strengthen the conditions used in [2]. The dual semigroup $\{\hat{P}_t,\ t \geq 0\}$ (see 2 , p. 191) is then defined on the state space E_∂, some of its properties will be reviewed and adduced. A process will then be constructed with the dual semigroup as its transition semigroup, which will be shown to be a Hunt process on E_∂. This process is in (strong) duality with the original Hunt Process in the sense of [1].

1. The Strengthened Hypotheses

Let $X = \{X_t,\ t \geq 0\}$ be a Hunt process with a locally compact, second countable state space E; and let $E_\partial = E \cup \{\partial\}$ be the one-point compactification of E. Let ξ be a Radon measure on $(E,\mathscr{E})$ where $\mathscr{E}$

*Research supported in part by NSF grant MCS83-01072.

is the Borel tribe on E. Let $\{P_t,\ t \geq 0\}$ be the transition semigroup, and U the potential kernel of X, such that

$$U(x,dy) = u(x,y)\xi(dy)\ . \tag{1}$$

The conditions on u and ξ are listed below for ready reference. Throughout this paper, the symbol K is reserved for a compact subset of E.

(i) $u(x,y) > 0$ for all $(x,y) \in E \times E$; $u(x,y) = +\infty$ if and only if $x = y \in E$;

(ii) for all $x \in E$: $u(x,\cdot)$ is extended continuous in E;

(iii) for all $y \in E$: $u(\cdot,y)$ is extended continuous in E;

(iv) for all $y \in E$: $\lim\limits_{x \to \partial} u(x,y) = 0$;

(v) there exists $x_o \in E$: $\lim\limits_{y \to \partial} u(x_o,y) = 0$;

(vi) for all K and for all $x \in E$: $\int_K u(x,y)\xi(dy) < \infty$;

(vii) for all K and for all $y \in E$: $\int_K \xi(dx)u(x,y) < \infty$;

(viii) ξ is an excessive measure.

Not all these conditions will be needed in all the results below. Let us compare them with those used in [2]. In that paper only conditions (i), (ii) and an alternative form of (vi) are imposed until §5, where (viii) and another condition implied by (iii) are added. Thus the new hypotheses are (iii), (iv), (v) and (vii). Let us remark that for the purposes of this paper, condition (iii) may be weakened as follows:

(iii') for all $y \in E$: $u(\cdot,y)$ is lower semicontinuous and bounded off a neighborhood of y.

Note that (ii) and (iv) together are equivalent to the continuity of

$u(\cdot,y)$ in E_∂ with $u(\partial,y) = 0$, for each $y \in E$. On the other hand, (iii) and (v) together imply only $u(x_o,\cdot)$ is continuous in E_∂ and $u(x_o,\partial) = 0$, for some $x_o \in E$. We shall set $u(x,\partial) = 0$ for each $x \in E$, without necessarily assuming the continuity of $u(x,\cdot)$ at ∂. Thus $u(x,y)$ is now defined for $(x,y) \in E_\partial \times E_\partial$ except at (∂,∂); and is equal to zero if $x = \partial$ or $y = \partial$.

The condition (vi), namely $U(\cdot,K)$ is a finite function, implies that there exists $h \in \mathcal{E}_+$ such that $0 < Uh < \infty$ on E, which in the presence of (iii) implies that

(2) $$\lim_{t \to \infty} P_t P_K 1 = 0 .$$

The last property is used as an assumption in [2], which is referred to as the transience of X. Let us remark that under (iii) or (iii'), $u(\cdot,y)$ is excessive for each y, hence the $\underline{u}(\cdot,y)$ which appears in [2] is simply $u(\cdot,y)$ here. Finally, let us recall the definition of a "potential" ([2], p. 169) as an excessive function s such that

(3) $$\lim_{K \uparrow E} P_{K^c} s = 0, \qquad \xi\text{-a.e.}$$

A "pure potential" is a potential such that

(4) $$\lim_{t \to \infty} P_t s = 0 \qquad \xi\text{-a.e.}$$

The set Q is defined ([2], p. 190) as follows:

(5) $$Q = \{y \in E: \ u(\cdot,y) \text{ is a pure potential}\}.$$

It is proved in [2] (Theorem 10, p. 195) that $Q \subset Z^c$. Hence in order to prove $Z = \emptyset$ it is sufficient to prove $Q = E$. The proof below does not use condition (v) or (viii).

THEOREM 1. For each $y \in E$, $u(\cdot,y)$ is a pure potential.

PROOF. As remarked above, $u(\cdot,y)$ is excessive. We have

$$(6) \qquad P_{K^c} u(x,y) = \int P_{K^c}(x,dz)u(z,y) .$$

As $P_{K^c}(x,dz)$ is concentrated on $\overline{K^c}$, while under condition (iv) $u(z,y)$ converges to zero as z leaves all compact subsets of E, it is clear that (3) is satisfied when $s = u(\cdot,y)$, indeed for all $x \in E$. Thus $u(\cdot,y)$ is a potential.

Fix a compact K and let V be open such that $y \in V \subset K^o$. For $t > 1$ we have

$$(7) \quad P_t u(x,y) \le \int_{t-1}^{t} P_s u(x,y)ds = \int_{t-1}^{t} \Big(\int_V + \int_{K-V} + \int_{E-K}\Big)P_s(x,dz)u(z,y)ds .$$

If $x \neq y$, then $u(x,y) < \infty$; we can choose V so that

$$\sup_{z \in V} u(x,z) = M_1 < \infty$$

by (ii); and

$$\int_V \xi(dz)u(z,y) \le \varepsilon$$

by (vii). It follows that

$$\int_{t-1}^{t} \int_V P_s(x,dz)u(z,y)ds \le \int_V U(x,dz)u(z,y) = \int_V u(x,z)u(z,y)\xi(dz) \le M_1\varepsilon.$$

Next, we have

$$\sup_{z \in E-V} u(z,y) \le M_2 < \infty$$

by (iii); hence

$$\int_{t-1}^{t} \int_{K-V} P_s(x,dz)u(z,y)ds \le \int_{K-V} P_{t-1} U(x,dz)u(z,y) \le M_2 P_{t-1} U(x,K).$$

Since $U(x,K) < \infty$, the last term above converges to zero as $t \to \infty$. Lastly we can choose K so that

$$\sup_{z \in E-K} u(z,y) \leq \varepsilon$$

by condition (iv). Hence

$$\int_{t-1}^{t} \int_{E-K} P_s(x,dz)u(z,y)ds \leq \varepsilon .$$

Combining the estimates above, we obtain if $x \neq y$:

$$\lim_{t \to \infty} P_t\, u(x,y) = 0 . \tag{8}$$

Thus (4) is true when $s = u(\cdot,y)$ with the only exceptional set $\{y\}$. Since $\xi(\{y\}) = 0$, we have proved $u(\cdot,y)$ is a pure potential.

Q.E.D.

COROLLARY. For any measure μ on $\mathcal{E}$, if $U\mu \not\equiv \infty$, then $U\mu$ is a pure potential.

PROOF. Recall that even under the more general conditions of [2] (Corollary on p. 171) we have $U\mu < \infty$, ξ-a.e. Let $U\mu(x) < \infty$, then $\mu(\{x\}) = 0$ because $u(x,x) = \infty$; and $P_t\, U\mu(x) < \infty$, $P_{K^c}\, U\mu(x) < \infty$ because $U\mu$ is excessive. It follws that

$$\infty > P_t\, U\mu(x) = \int_{E-\{x\}} P_t\, u(x,y)\mu(dy) \downarrow 0$$

as $t \to \infty$, by (8); and

$$\infty > P_{K^c} U\mu(x) = \int_E P_{K^c}\, u(x,y)\mu(dy) \downarrow 0$$

as $K \uparrow E$, by the argument under (6).

We state some of the consequences of the result "$Z = \emptyset$", which we have just proved.

(a) For each $y \in E$, and open neighborhood G of y, we have

$$u(x,y) = P_G\, u(x,y), \qquad x \in E, \quad y \in G.$$

This is actually equivalent to "$Z = \emptyset$", see p. 173 of [2]; but we shall not use this directly.

(b) For any two measures μ and ν on $\mathcal{E}$, if $U\mu = U\nu \not\equiv \infty$ on $\mathcal{E}$ then $\mu = \nu$, see p. 186 of [2].

(c) Riesz representation. If f is an excessive function $\not\equiv \infty$, then there exists a unique Radon measure μ and a harmonic function such that

$$f = U\mu + h. \tag{9}$$

It is proved in Theorem 6, p. 187 of [2], under more general conditions, that we have such a representation with $\mu(Z) = 0$, and that $Q \subset Z^c$. But it does not follow that $U\mu$ is a potential since μ may change $Q^c - Z$. Under the present conditions, $U\mu$ is a potential by the Corollary to Theorem 1. In [2], the uniqueness of μ is proved only for the class of μ with $\mu(Z) = 0$, so that the possibility of another representation like (9) but with $\mu(Z) > 0$ cannot be ruled out. The absence of Z ensures the uniqueness of representation.

Finally, we will record a simplified form of the principal convergence theorem in [2]. This combines Theorem 2 and its continuation on p. 176 and p. 180 of [2], and corrects a slip there.

THEOREM 2. Suppose $Z = \emptyset$. Let $\{\mu_n\}$ be a sequence of measures on $\mathcal{E}$ satisfying the conditions:

(a) for all n: $U\mu_n \le \sigma$, where σ is a potential;

(b) $\lim\limits_n U\mu_n = s$, where s is an excessive function $\not\equiv \infty$.

Then $\{\mu_n\}$ converges vaguely to a Radon measure μ, and $s = U\mu$.

The word "diffuse" in the original statement of Theorem 2 should be deleted. The only place this superfluous (and inconvenient) condition was apparently used is in (2) on p. 177 of [2]. But since $u(x,x) = \infty$, $U\mu_n(x) < \infty$ implies $\mu_n(\{x\}) = 0$, which is what we need there. The vague convergence of the entire sequence $\{\mu_n\}$ follows from the unique determination of any subsequential vague limit, on account of consequence (b) above.

2. The Dual Semigroup

Recall the definition of the dual semigroup $\{\hat{P}_t, t \ge 0\}$ on p. 190 of [2]*. For each $y \in E$ and $t \ge 0$, $P_t u(\cdot,y)$ is a pure potential by Theorem 1 above. Hence by the Riesz theorem there is a unique measure μ such that $P_t u(\cdot,y) = U\mu$. This measure is denoted by $\hat{P}_t(\cdot,y)$. Thus we have for all x and y:

$$(1) \qquad P_t u(x,y) = \int P_t(x,dz)u(z,y) = \int u(x,z)\hat{P}_t(dz,y).$$

Let $g \in b\mathcal{E}_+$ with compact support, so that $Ug < \infty$ by condition (vi). We have by (1)

$$(2) \qquad \int u(x,z)\int\hat{P}_t(dz,y)g(y)\xi(dy) = \int P_t u(x,y)g(y)\xi(dy) = P_t Ug(x)$$
$$= UP_t g(x) = \int u(x,z)P_t g(z)\xi(dz).$$

It follows from uniqueness of the Riesz representation that

$$(3) \qquad \int\hat{P}_t(dz,y)g(y)\xi(dy) = P_t g(z)\xi(dz)$$

*It can be verified that $\{\hat{P}_t\}$ is Borelian.

as measures. For any $f \in b\mathcal{E}_+$, if we integrate f with respect to the two measures in (3), we obtain

$$(4) \qquad \int (f\hat{P}_t)g \, d\xi = \int f(P_t g) d\xi .$$

This is the duality relation between (P_t) and $(\hat{P}_t)$.

Since $Q = E$, the following results are proved in Theorem 7, p. 191, Theorem 8, p. 193, and (31), p. 195 of [2].

THEOREM 3. $\{\hat{P}_t, t \geq 0\}$ is a submarkovian semigroup on $E \times E$. For each y, $\hat{P}_t(\cdot,y)$ is weakly right continuous in $t \geq 0$, and $\hat{P}_o(\cdot,y) = \delta_y$. Furthermore, we have

$$(5) \qquad \hat{U}(dx,y) = \xi(dx)u(x,y) .$$

The weak right continuity is a consequence of vague right continuity, the semigroup property, and the weak right continuity at $t = 0$. The proof depends on Theorem 2 in §1. Under condition (vii), $\hat{U}(K,y) < \infty$ for each $y \in E$ and compact $K \subset E$.

We extend $(\hat{P}_t)$ to E_∂ to be strictly Markovian in the usual way. Namely, we set $\hat{P}_t(\{\partial\},\partial) = 1$ for $t \geq 0$; and $\hat{P}_t(\{\partial\},y) = 1 - \hat{P}_t(E,y)$ for $y \in E$, $t \geq 0$. We denote this extension by $(\hat{P}_t)$ without changing the notation.

3. Coexcessive Functions

A function on E is coexcessive iff f is excessive with respect to $(\hat{P}_t)$ on E. If we put $f(\partial) = 0$, this can be extended to E_∂. We shall do so in the sequel. Since $\hat{U}(K,\cdot) < \infty$ for each compact $K \subset E$, it follows from a general theorem ([3], p. 86) that: if f is coexcessive, then there exist $g_n \in b\mathcal{E}_+$ such that

$$(1) \qquad f = \lim_{n \to \infty} \uparrow g_n \hat{U} .$$

In fact we may choose $g_n \leq n^2$ and $g_n \hat{U} \leq n$. By condition (ii) and Fatou's lemma, $g_n \hat{U}$ is l.s.c. (lower semicontinuous). Hence each coexcessive function is l.s.c. by (1). It is a consequence of (1) and $u > 0$ that a coexcessive function $\not\equiv 0$ is strictly positive on E (cf. Proposition 9 on p. 171 of [2]). If f_1 and f_2 are both coexcessive, then $f_1 \wedge f_2$ is superaveraging with respect to $(\hat{P}_t)$, and l.s.c. Since $(\hat{P}_t)$ is weakly right continuous, it follows that $f_1 \wedge f_2$ is coexcessive. Similarly if f is coexcessive and c is a constant ≥ 0, then $f \wedge c$ is coexcessive. For each $x \in E$, $u(x,\cdot)$ is coexcessive by (1) of §2, because for each $y \in E$, $u(\cdot,y)$ is excessive.

Let $C_o(E_\partial)$ denote the class of continuous functions on E_∂ which vanish at ∂. We define

S = the class of coexcessive functions which belong to $C_o(E_\partial)$;

$L = S - S$.

By condition (v), $u(x_o,\cdot) \in S$. Indeed, we may replace that condition by the following:

(v') there exists a member of S which is not identically zero.

It will be seen that all the arguments below remain valid if we replace $u(x_o,\cdot)$ by any member $\not\equiv 0$ of S.

It seems of interest to examine the significance of condition (v). This has to do with the lifetime $\hat{\zeta}$ of the dual process. Let

$$\lambda(y) = \hat{P}^y\{\hat{\zeta} = \infty\} , \qquad y \in E_\partial .$$

Then λ is coexcessive, and it follows as in Proposition 9, p. 171 of [2] that if $\lambda \not\equiv 0$ in E, then $\lambda > 0$ in E. The following result can be proved in a way similar to that of Theorem 1.

PROPOSITION 4. Suppose $\lambda > 0$ in E. If for any $x \in E$, the function $u(x,\cdot)$ is continuous in E_∂, then $u(x,\partial) = 0$.

A dual proposition for $u(\cdot,y)$ is also true. Thus, suppose $P^x\{\zeta = \infty\} > 0$ for some (hence all) $x \in E$, and $u(\cdot,y)$ is purely excessive. If $u(\cdot,y)$ is continuous at ∂, then $u(\partial,y) = 0$.

From now on all the conditions in §1 will be used.

THEOREM 5. L is dense in $C_o(E_\partial)$ endowed with the sup-norm.

PROOF. S is a cone which is closed under the minimum operation, and also under truncation by a constant $c \geq 0$, as reviewed above. Let $x_1 \neq x_2$, both in E. Then since $u > 0$ on $E \times E$, there exists a constant $A > 0$ such that $Au(x_o,x_1) > u(x_1,x_2)$, where x_o is the point in condition (v). Put $\varphi(y) = Au(x_o,y) \wedge u(x_1,y)$. Then $\varphi \in S$ and $\varphi(x_1) \neq \varphi(x_2)$. Next put $\varphi(y) = u(x_o,y) \wedge 1$. Then $\varphi \in S$ and $\varphi(x) \neq \varphi(\partial)$ for any $x \in E$. Thus S separates the points of E_∂. Hence so does L. Let $f_1 \in L$, $f_2 \in L$, then $f_1 - f_2 = g_1 - g_2$ where $g_1 \in S$, $g_2 \in S$. Hence $g_1 \wedge g_2 \in S$, $g_1 + g_2 \in S$, $|f_1 - f_2| = |g_1 - g_2| = g_1 + g_2 - 2(g_1 \wedge g_2) \in L$, $f_1 \wedge f_2 = \frac{1}{2}\{f_1 + f_2 - |f_1 - f_2|\} \in L$, $f_1 \vee f_2 = f_1 + f_2 - (f_1 \wedge f_2) \in L$. Therefore L is a lattice. It is trivial that L is also a vector space.

Let K be a compact subset of E and let $L(K)$ denote the class of functions of L restricted to K. Let $\inf_{y \in K} u(x_o,y) = b > 0$. For any constant $c \geq 0$ put $\varphi(y) = \frac{c}{b} u(x_o,y) \wedge c$. Then $\varphi \in S$ and $\varphi = c$ on K. Therefore $L(K)$ contains all constants $c \geq 0$. It is also a lattice and a vector space. By a form of the Stone-Weierstrass theorem (see e.g. [5], p. 172), $L(K)$ is dense in $C(K)$ = the class of continuous functions on K.

Let $f \geq 0$ and have compact support $K \subset E$. For any $\varepsilon > 0$ there exists $g \in L$ such that $|f-g| \leq \varepsilon$ on K. Since $f \geq 0$, it is trivial that $|f-g^+| \leq \varepsilon$ on K. Since $g^+ \in C_o(E_\partial)$, there exists a compact K_1 such that $K \subset K_1 \subset E$ and $g^+ \leq \varepsilon$ on $E_\partial - K_1$. Hence, as before, there exists $h \in L$ such that $|f-h^+| \leq \varepsilon$ on K_1. Put $\varphi = g^+ \wedge h^+$; then $\varphi \in L$. We have $|f-\varphi| \leq \varepsilon$ on K; $|f-\varphi| = \varphi \leq h^+ \leq \varepsilon$ on $K_1 - K$; $|f-\varphi| = \varphi \leq g^+ \leq \varepsilon$ on $E_\partial - K_1$. Hence $|f-\varphi| \leq \varepsilon$ on E_∂. Since L is a vector space, it follows that L is dense in the class of continuous functions having compact support in E; hence it is also dense in $C_o(E_\partial)$.

Q.E.D.

4. The Dual Process

It is well known how to construct a Markov process on E_∂ with $(\hat{P}_t)$ as its transition semigroup. Let $\{Y_t,\ t \geq 0\}$ be such a process, and $\mathcal{G}^o_t = \sigma(Y_s,\ 0 \leq s \leq t)$, $\mathcal{G}^o = \bigvee_{0 \leq t < \infty} \mathcal{G}^o_t$ its natural filtration. We can define, for each $y \in E_\partial$, the probability $\hat{P}^y$ and expectation $\hat{E}^y$ on $\mathcal{G}^o$, associated with the process in the usual way. We proceed to show that there is a version of this process whose paths are almost surely (a.s.) right continuous in $[0,\infty)$ and have left limits in $(0,\infty]$. Although $(\hat{P}_t)$ is in general not a Feller semigroup, standard methods developed for the latter case in Chapter 2, §§2-4 of [3] can be adapted to the present situation with easy modifications. We shall indicate the main steps below.

(A) The process $\{Y_t\}$ is stochastically right continuous.

This follows from the weak (or just vague) right continuity of $\hat{P}_t$; see p. 50 of [3].

(B) Let R be the set of rational numbers in $[0,\infty)$. Then a.s. the restriction of the sample function $t \to Y_t$ to R has right limits

in $[0,\infty)$ and left limits in $(0,\infty]$ everywhere.

Since $S \subset C_o(E_\partial)$ and the latter is a separably metrizable space, there is a countable subset D which is dense in S (with respect to the sup-norm of $C_o(E_\partial)$). If we use this set D instead of the $\mathbb{D}$ on p. 53 of [3], the same argument there works. Note that left limits exist at ∞, because a positive supermartingale has such a limit a.s.

(C) There is a version of (Y_t) whose sample functions are a.s. right continuous in $[0,\infty)$ and have left limits in $(0,\infty]$.

The argument is exactly the same as on p. 54 of [3]. From now on we shall use this version and refer to it as the dual process. Its lifetime will be denoted by $\zeta = T_{\{\partial\}}$. Then we have a.s.

$$Y_{t-} \in E \qquad \text{and} \qquad Y_t \in E \qquad \text{for} \quad 0 \leq t < \zeta .$$

The fact $\zeta > 0$ a.s. follows from $\hat{P}^y\{Y_o = y\} = \hat{P}_o(\{y\},y) = 1$ and right continuity. The rest is proved exactly as on pp. 54-55 of [3], if we use the member of S postulated in condition (v) or (v'), instead of the function $U^1\varphi$ there.

(D) For any coexcessive function f, the function $t \to f(X_t)$ is a.s. right continuous in $[0,\infty)$, and has left limits in $(0,\infty]$.

PROOF. Let g_n be as in (1) of §3, and let $K_j \uparrow E$, where each K_j is compact. Put

$$(1) \qquad \varphi_{nj}(y) = \int \xi(dx) g_n(x) 1_{K_j}(x)[u(x,y) \wedge ju(x_o,y) \wedge j] .$$

It follows from conditions (ii), (v) and (vii) that $\varphi_{nj} \in C_o(E_\partial)$. Hence $t \to \varphi_{nj}(Y_t)$ is an a.s. right continuous positive supermartingale. Since $g_n\hat{U} = \lim_{j\to\infty} \uparrow \varphi_{nj}$, the same is true of $t \to g_n\hat{U}(Y_t)$ by a theorem of Meyer's (see Theorem 5, §1.4 of [3], p. 32). Since $f = \lim_{n\to\infty} \uparrow g_n\hat{U}$,

another application of the theorem establishes the right continuity of $t \to f(Y_t)$. The existence of left limits is then a consequence.

Meyer [4] proved that for a right continuous homogeneous Markov process, right continuity of all α-excessive functions along the paths is equivalent to the strong Markov property of the process. The proof given below follows his argument in one direction, but uses only 0-excessive functions.

(E) The dual process has the strong Markov property.

PROOF. Let T be optional with respect to $\{G^o_{t+}, t \geq 0\}$, and $T_n = 2^{-n}[2^n T + 1]$. Let f be positive continuous with compact support so that $f\hat{U} < \infty$ by (vii). We have by the simple Markov property, for each $y \in E$, $s \geq 0$:

$$\hat{E}^y\{f(Y_{T_n+s})|G^o_{T_n}\} = f\hat{P}_s(Y_{T_n}) .$$

Integrating over s, we obtain

$$(2) \qquad \hat{E}^y\{\int_{T_n+t}^{\infty} f(Y_s)ds|G^o_{T_n}\} = \int_t^{\infty} \hat{E}^y\{f(Y_{T_n+s})|G^o_{T_n}\}ds = \int_t^{\infty} f\hat{P}_s(Y_{T_n})ds$$

$$= f\hat{U}\hat{P}_t(Y_{T_n}) .$$

Since $\hat{E}^y\{\int_0^{\infty} f(Y_s)ds\} = f\hat{U}(y) < \infty$, as $n \to \infty$ the first member in (2) converges to

$$\hat{E}^y\{\int_{T+t}^{\infty} f(Y_s)ds|G^o_{T+}\} = \int_t^{\infty} \hat{E}^y\{f(Y_{T+s})|G^o_{T+}\}$$

by a well-known dominated convergence theorem for conditional expectations (see [6], Theorem 9.4.8). Since $f\hat{U}\hat{P}_t$ is coexcessive, the last member in (2) converges to $f\hat{U}\hat{P}_t(Y_T)$ as $n \to \infty$, by (D). The result is

$$(3) \qquad \int_t^{\infty} \hat{E}^y\{f(Y_{T+s})|G^o_{T+}\}ds = \int_t^{\infty} f\hat{P}_s(Y_T)ds .$$

The integrand in the left member of (3) is right continuous in s; so is that in the right member because $\hat{P}_s$ is weakly right continuous in $s \geq 0$. Hence it follows by differentiation of (3) that

$$\hat{E}^y\{f(Y_{T+t})|G^o_{T+}\} = f\hat{P}_t(Y_T) \tag{4}$$

for all $t \geq 0$. This implies the strong Markov property of the dual process.

(F) The dual process has the moderate Markov property.

PROOF. Let T be predictable with respect to $\{G^o_{t+}\}$, and let $\{T_n\}$ announce T. For each n, we have by (E), for any $f \in C_o(E_\partial)$:

$$\hat{E}^y\{f(Y_{T_n+t})|G^o_{T_n+}\} = f\hat{P}_t(Y_{T_n}) \ . \tag{5}$$

If $f \in S$, then $f\hat{P}_t$ is l.s.c. Letting $n \to \infty$ so that $Y_{T_n+t} \to Y_{T+t-}$ for both t and $t = 0$, we obtain

$$\hat{E}^y\{f(Y_{T+t-})|G^o_{T-}\} \geq f\hat{P}_t(Y_{T-}) \tag{6}$$

where $Y_{o-} = Y_o$. Now we let $t \downarrow\downarrow 0$, then $Y_{T+t-} \to Y_{T+} = Y_T$, while $f\hat{P}_t \to f$ by coexcessivity. It follows that

$$\hat{E}^y\{f(Y_T)|G^o_{T-}\} \geq f(Y_{T-}) \ . \tag{7}$$

Since $\{f(Y_t),\ G^o_{t+}\ ,\ t \geq 0\}$ is a positive supermartingale, we have by the stopping theorem, for each n:

$$\hat{E}^y\{f(Y_T)|G^o_{T_n+}\} \leq f(Y_{T_n}) \ . \tag{8}$$

When $n \to \infty$ in (8), the result is an inequality which is reverse to (7). Therefore, (7) holds with equality for all $f \in S$, hence for all $f \in C_o(E_\partial)$ by Theorem 5. This implies $\hat{E}^y\{Y_T = Y_{T-}\} = 1$ by Lemma 1 on p. 66 of [3].

The quasi left continuity of the dual process follows by a general argument given on p. 70 of [3].

(G) Augmentation of $\{G_t^o, t \geq 0\}$.

Exactly as detailed on pp. 61-62 of [3], the natural filtration for the dual process can be augmented so that the new filtration $\{G_t, t \geq 0\}$ is right continuous. The strong and moderate Markov properties proved above are then valid for optional and predictable times with respect to the augmented filtration.

With these technical ramifications, we conclude that the dual process $\{Y_t, G_t; t \geq 0\}$ is a Hunt process with the semigroup $\{\hat{P}_t, t \geq 0\}$.

(H) Strong duality.

The duality relation in (4) of §2 implies the following. For each $\alpha \geq 0$, $f \in \mathcal{E}_+$, $g \in \mathcal{E}_+$, we have

$$\int (f\hat{U}^\alpha) g \, d\xi = \int f(U^\alpha g) d\xi ,$$

where

$$U^\alpha g = \int_0^\infty e^{-\alpha t} P_t g \, dt , \qquad f\hat{U}^\alpha = \int_0^\infty e^{-\alpha t} f\hat{P}_t \, dt .$$

Moreover, since

$$U(x,dy) = u(x,y)\xi(dy), \qquad \hat{U}(dx,y) = \xi(dx)u(x,y) ,$$

both $U(x,\cdot)$ and $\hat{U}(\cdot,y)$ are absolutely continuous with respect to ξ for each x and y; hence so are $U^\alpha(x,\cdot)$ and $\hat{U}^\alpha(\cdot,y)$, for $\alpha > 0$. Thus the hypotheses (referred to as those of "strong duality") in Theorem 1.4, §6.1 of [1] are satisfied. Therefore, all consequences of these hypotheses developed there apply to the two transient Hunt processes X and Y in this paper.

REFERENCES

1. R.M. BLUMENTHAL and R.K. GETOOR. *Markov Processes and Potential Theory.* Academic Press, New York, 1968.
2. K.L. CHUNG and K.M. RAO. A new setting for potential theory. *Ann. Inst. Fourier 30* (1980), 167-198.
3. K.L. CHUNG. *Lectures from Markov Processes to Brownian Motion.* Springer-Verlag, Berlin, 1982.
4. P.A. MEYER. *Processes de Markov.* Lecture Notes in Mathematics 26, Springer-Verlag, Berlin, 1967.
5. H.L. ROYDEN. *Real Analysis, 2nd Ed.* Macmillan Co., New York, 1968.
6. K.L. CHUNG. *A Course in Probability Theory, 2nd Ed.* Academic Press, New York, 1974.

K. L. Chung
Ming Liao
Department of Mathematics
Stanford University
Stanford, CA 94305

K. M. Rao
Department of Mathematics
University of Florida
302 Walker Hall
Gainesville, FL 32611

Reprinted from *Séminaire de Probabilités*, XIX, 1983/1984, 496–503,
Lecture Notes in Mathematics, 1123, Springer-Verlag, Berlin (1985).
With kind permission of Springer Science and Business Media.

THE GAUGE AND CONDITIONAL GAUGE THEOREM

K. L. Chung*

Let $\{X_t, t \geq 0\}$ be the Brownian motion in R^d, $d \geq 1$. Let D be a bounded domain in R^d, $\bar{D}$ its closure, ∂D its boundary; and let q be a Borel function defined in R^d and satisfying the following condition:

$$(1) \qquad \lim_{t \downarrow 0} \sup_{x \in D} E^x\{\int_0^t 1_D|q|(X_s)ds\} = 0$$

where 1_D is the indicator of D. Such a function is said to belong to the Kato class K_d. The equivalent condition (1) is given by Aizenman and Simon [1].

The *gauge* for (D,q) is defined to be the function u on $\bar{D}$ below:

$$(2) \qquad u(x) = E^x\{\exp(\int_0^{\tau_D} q(X_s)ds)\} \quad .$$

From here on we write for abbreviation:

$$(3) \qquad e_q(t) = \exp(\int_0^t q(X_s)ds) \; .$$

For a domain D with $m(D) < \infty$ (where m denotes the Lebesgue measure), without any regularity hypothesis on ∂D, and a bounded q, we proved the following theorem in [3].

The Gauge Theorem. *If* $u(\cdot) \not\equiv \infty$ *in* D, *then* u *is bounded in* $\bar{D}$.

*Research supported in part by NSF grant MCS 83-01072 at Stanford University.

Actually, if ∂D is regular in the sense of the Dirichlet problem, then u is continuous and strictly positive in $\bar{D}$. However, in this note we shall concentrate on the main thing, as stated above. Zhao [6] extended the theorem to $q \in K_d$ for a bounded domain in R^d, $d \geq 3$; he also did the case $d = 2$ in yet unpublished notes. For $d = 1$ and D a half-line, see [2]. Prior to Zhao's work, Falkner extended the theorem in another direction by considering the *conditional gauge* for (D,q) as follows:

$$(4) \qquad u(x,z) = E_z^x\{e_q(\tau_D)\}, \qquad (x,z) \in D \times \partial D;$$

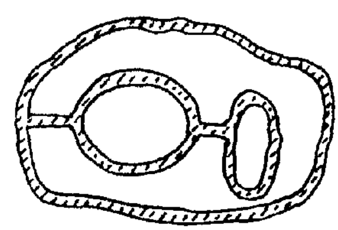
where E_z^x is the expectation associated with the Brownian motion killed outside D, starting from x and conditioned to converge to z (at its life-time τ_D). For a class of bounded domains including those with C^2 boundary, and bounded q, he proved the following theorem in [5].

<u>Conditional Gauge Theorem</u>. *If* $u(\cdot,\cdot) \not\equiv \infty$ *in* $D \times \partial D$, *then it is bounded there. This is the case if and only if* $u(\cdot) \not\equiv \infty$ *in* D, *as in the gauge theorem.*

I gave a simpler proof of Falkner's theorem in [4]. Subsequently Zhao [7] proved that if $u(\cdot) \not\equiv \infty$, then $u(\cdot,\cdot) \not\equiv \infty$, for bounded C^2 domains. He has since proved the conditional gauge theorem as stated above for bounded $C^{1,1}$ domains. In this note I shall show that the conditional gauge theorem actually follows in a general way and rather quickly from the gauge theorem.

The basic probabilistic argument turns out to be an old one in [2] (see the proof of Theorem 1 there), easily adapted to the multi-dimensional case. The sole difficulty encountered in extending the class of bounded q to the class K_d is contained in Lemma 1 below.

We begin by setting up the framework of the probabilistic argument involving a sequence of hitting times. Let D_1 and D_2 be subdomains of D such that $\bar{D}_1 \subset D_2$, $\bar{D}_2 \subset D$, and $C \triangleq D - \bar{D}_1$ is connected and $m(C) < \varepsilon$ for an arbitrary $\varepsilon > 0$. This is possible if ∂D is Lipschitzian for instance. For then each connected component of $R^d - \bar{D}$ must contain a ball of fixed size, hence there are at most a finite number of "holes" inside the outer boundary of D. Since D is connected, it is easy to see how to construct D_1 and D_2 as desired. A picture illustrates the result. I am indebted to Falkner for alerting me to the necessity of making C connected.



(The shaded portion represents C.)

Lemma 1. *If* ∂D *is sufficiently smooth, then for any given* $\varepsilon > 0$, *there exists* $\delta(\varepsilon)$ *such that if the* C *described above has* $m(C) \leqslant \delta(\varepsilon)$, *then*

499

(5) $$\sup_{\substack{x \in C \\ z \in \partial D}} E_z^x\{\int_0^{\tau_C} |q|(X_t)dt\} \leq \varepsilon \ ;$$

(6) $$\sup_{\substack{x \in C \\ z \in \partial D}} E_z^x\{e_q(\tau_C)\} \leq \frac{1}{1-\varepsilon} \ .$$

In [7], Zhao proved that C^2 boundary is sufficient for the lemma to hold; more recently he has improved this result to require only $C^{1,1}$ boundary. In this connection it should be mentioned that the gauge theorem for an arbitrary bounded domain D, and $q \in K_d$, follows quickly from an easier analogue of (5) for a small ball B, as follows:

(7) $$\sup_{\substack{x \in B \\ z \in \partial B}} E_z^x\{\int_0^{\tau_B} |q|(X_t)dt\} \leq \varepsilon \ .$$

This was proved in Zhao [6]. The deduction of (6) from (5) is standard Markovian calculation.

Lemma 2 is a strengthened form of an argument I have indicated elsewhere (see [5], Remark 2.13). The constants $a_1, a_2, \cdots$ below are strictly positive, depending only on D_1, D_2 and D. We assume ∂D to be Lipschitzian below.

<u>Lemma 2</u>. *For all* $y \in \partial D_2$ *and* $z \in \partial D$, *we have*

(8) $$a_1 \leq E_z^y\{\tau_C = \tau_D;\ e_q(\tau_D)\} \leq a_2 \ .$$

Proof: Recall that

(9) $$P_z^y\{\tau_C < \tau_D\} = \frac{f(y,z)}{K(y,z)}$$

where K is the Poisson kernel for D, and

$$f(y,z) = E^y\{\tau_C < \tau_D;\ K(X(\tau_C),z)\} .$$

For each $y \in \partial D_2$, $f(y,\cdot)$ is continuous on ∂D, because on $\{\tau_C < \tau_D\}$ we have $X(\tau_C) \in \partial D_1$ almost surely, and K is bounded continuous in $\partial D_1 \times \partial D$. For each $z \in \partial D$, $f(\cdot,z)$ is harmonic in C. Hence f is continuous on $\partial D_2 \times \partial D$ since ∂D_2 and ∂D are disjoint closed sets. It follows that the function of (y,z) in (9) is continuous and positive on $\partial D_2 \times \partial D$. The function $K(\cdot,z) - f(\cdot,z)$ is harmonic in C and unbounded in the neighborhood of z, because K is unbounded while f is bounded. Hence it is strictly positive in C by harmonicity, because C is connected and $z \in \partial C$. Therefore we have by continuity

(10) $$b \triangleq \inf_{\substack{y \in \partial D_2 \\ z \in \partial D}} P_z^y\{\tau_C = \tau_D\} > 0 .$$

Now it follows by Jensen's inequality and (15) that for $(y,z) \in \partial D_2 \times \partial D$:

(11) $$\begin{aligned} E_z^y\{e_q(\tau_D)\,|\,\tau_C = \tau_D\} &\geq E_z^y\{e_{-|q|}(\tau_D)\,|\,\tau_C = \tau_D\} \\ &\geq \exp\{-E_z^y[\int_0^{\tau_D} |q|(X_t)dt\,|\,\tau_C = \tau_D]\} \\ &\geq \exp\{-\frac{1}{b}\int_0^{\tau_C} |q|(X_t)dt\} \geq e^{-\varepsilon/b} . \end{aligned}$$

501

Combining (10), (11) and (16), we have proved (8) with $a_1 = b\, e^{-\varepsilon/b}$, $a_2 = \frac{1}{1-\varepsilon}$.

We are ready to prove the conditional gauge theorem for a bounded Lipschitzian domain for which the conclusions of Lemma 1 hold true, thus at least when ∂D belongs to $C^{1,1}$. Put $T_o \equiv 0$, and for $n \geq 1$:

$$T_{2n-1} = T_{2n-2} + \tau_{D_2} \circ \theta_{T_{2n-2}},$$

$$T_{2n} = T_{2n-1} + \tau_C \circ \theta_{T_{2n-1}}.$$

For any $(x,z) \in D \times \partial D$, we have $P_z^x\{\tau_D < \infty\} = 1$. This nontrivial result has recently been proved by M. Cranston for a bounded Lipschitzian domain; for a bounded C^1-domain D it follows from the fact that $K(\cdot,z)$ is integrable over D, by a remark communicated to me by Kenig. It follows that for some $n \geq 1$, $X(T_{2n}) \in \partial D$. Therefore we have by the strong Markov property of the conditioned process:

$$(12)\qquad E_z^x\{e_q(\tau_D)\} = \sum_{n=1}^{\infty} E_z^x\{T_{2n} = \tau_D;\ e_q(\tau_D)\}$$

$$= \sum_{n=1}^{\infty} E_z^x\{T_{2n-2} < \tau_D;\ e_q(T_{2n-1}) E_z^{X(T_{2n-1})}[\tau_C = \tau_D; e_q(\tau_D)]\}.$$

Observe that $\partial C = \partial D_1 \cup \partial D$. On the set $\{T_{2n-2} < \tau_D\}$, $X(T_{2n-1}) \in \partial D_2$. Hence by Lemma 2

$$(13)\qquad a_1 \sum_{n=1}^{\infty} E_z^x\{T_{2n-2} < \tau_D; e_q(T_{2n-1})\} \leq E_z^x\{e_q(\tau_D)\}$$

$$\leq a_2 \sum_{n=1}^{\infty} E_z^x\{T_{2n-2} < \tau_D; e_q(T_{2n-1})\}.$$

The general term in the series above is explicitly:

$$(14) \qquad \frac{1}{K(x,z)} E^x\{T_{2n-2} < \tau_D;\ e_q(T_{2n-1})K(X(T_{2n-1}),z)\} .$$

Since K is continuous and strictly positive on $\overline{D}_2 \times \partial D$, we have for (x,z) and (x',z') in $\overline{D}_2 \times \partial D$, almost surely

$$(15) \qquad a_3 \frac{K(X(T_{2n-1}),z')}{K(x',z')} \leqslant \frac{K(X(T_{2n-1}),z)}{K(x,z)} \leqslant a_4 \frac{K(X(T_{2n-1}),z')}{K(x',z')}$$

where a_3 and a_4 depend only on $\overline{D}_2$ and D. It follows from (13), (14) and (15) that

$$(16) \qquad \sup_{x \in \overline{D}_2} \sup_{z \in \partial D} u(x,z) \leqslant \frac{a_2 a_4}{a_1 a_3} \inf_{x \in \overline{D}_2} \inf_{z \in \partial D} u(x,z) .$$

Since $u(x)$ is a probability average of $u(x,z)$ over $z \in \partial D$, we have

$$(17) \qquad \inf_{z \in \partial D} u(x,z) \leqslant u(x) \leqslant \sup_{z \in \partial D} u(x,z) .$$

Now by hypothesis of the theorem, there exists $(x_o,z_o) \in D \times \partial D$ such that $u(x_o,z_o) < \infty$. Without loss of generality we may suppose $x_o \in D_2$. Hence by (16)

$$(18) \qquad \sup_{z \in \partial D} u(x_o,z) \leqslant \frac{a_2 a_4}{a_1 a_3} u(x_o,z_o) < \infty .$$

Next by (17), $u(x_o) < \infty$. Hence by the gauge theorem, $\sup_{x \in \overline{D}} u(x) < \infty$. It follows then by (16) and (17) that

$$(19) \qquad \sup_{x \in \overline{D}_2} \sup_{z \in \partial D} u(x,z) < \infty .$$

503

For $x \in D - \overline{D}_2$, we use the old argument in [3] adapted to the conditioned process, as follows:

$$u(x,z) = E_z^x\{\tau_C = \tau_D;\ e_q(\tau_C)\} + E_z^x\{\tau_C < \tau_D;\ u(X(\tau_C),z)\}$$

$$\leqslant \frac{1}{1-\varepsilon} + \sup_{x \in \overline{D}_1} \ \sup_{z \in \partial D} u(x,z) < \infty .$$

This establishes the first assertion of the conditional gauge theorem. The second assertion has also been proved between the lines above.

Remark : Conditional gauge theorem is also true for a bounded C^1 domain, and bounded q, using a hard inequality of Kenig's to prove lemma 1.

References

[1] Aizenman, N., Simon, B.: Brownian motion and Harnack inequality for Schrödinger operators, Comm. Pure Appl. Math. 35 (1982), 209-273.

[2] Chung, K.L.: On stopped Feynman-Kac functionals, Séminaire de Probabilités XIV, 1978/79, Lecture Notes in Mathematics No. 784, Springer-Verlag.

[3] Chung, K.L., Rao, K.M.: Feynman-Kac functional and Schrödinger equation, Seminar on Stochastic Processes 1, 1-29, Birkhäuser 1981.

[4] Chung, K.L.: Conditional gauges, Seminar on Stochastic Processes 3, 1983.

[5] Falkner, N.: Feynman-Kac functionals and positive solutions of $\frac{1}{2}\Delta u+qu = 0$, Z. Wahrsch. Verw. Gebiete 65 (1983), 19-33.

[6] Zhao, Z.: Conditional gauge with unbounded potential, Z. Wahrsch Verw. Gebiete 65 (1983), 13-18.

[7] Zhao, Z.: Uniform boundedness of conditional gauge and Schrödinger equations, Comm. Math. Phys 93 (1984), 19-31.

Seminar on Stochastic Processes, 1984
Birkhäuser, Boston, 1986

GAUGE THEOREM FOR THE NEUMANN PROBLEM

by

K.L. CHUNG* and PEI HSU*

Let D be a bounded domain in $\mathbb{R}^d$ and let $(\Delta/2 + q)u = 0$ be Schrödinger's equation on D. The Dirichlet problem for the equation was studied first in [2] for bounded q and then in [1] and [4] for $q \in K_d$ (see below for definition). The gauge function for the Dirichlet problem is defined in [2] as

$$(1) \qquad G(x) = E^x[\exp(\int_0^{\tau_D} q(B_s)\, ds)],$$

where $B = \{B_t, t \geq 0\}$ is the standard Brownian motion on $\mathbb{R}^d$ and τ_D is the first exit time of D. One striking property of the gauge function proved in [2] and [4] is the following

THEOREM 1. *If* G *is not identically infinite, then it is bounded on* $\bar{D}$.

The gauge function plays a key role in the solution of the Dirichlet problem; see the references mentioned above.

In this paper, we define a gauge function which plays a role in the Neumann problem similar to that of the gauge in the Dirichlet

* Research supported in part by NSF grant MCS83-01072 at Stanford University.

problem. It turns out that this new function also has the property stated in Theorem 1.

To define this gauge function, let us start with the definition of the class K_d.

DEFINITION. A measurable function q is in the class K_d iff

$$(2) \qquad \lim_{\alpha\to 0} \sup_{x\in R^d} \int_{|x-y|\le\alpha} G_d(x,y)\,|q|(y)\,dy = 0,$$

where G_d is the fundamental solution of Laplace's equation in R^d, namely

$$G_d(x,y) = \begin{cases} |x-y|, & \text{if } d = 1; \\ \log|x-y|^{-1}, & \text{if } d = 2; \\ |x-y|^{-2+d}, & \text{if } d \ge 3. \end{cases}$$

It is proved in [1] that condition (2) is equivalent to the condition

$$(3) \qquad \lim_{t\to 0} \sup_{x\in R^d} \int_0^t \int_{R^d} |q|(y)\Gamma(s,x,y)\,dy\,ds = 0,$$

where $\Gamma(t,x,y)$ is the transition density function of the standard Brownian motion on R^d:

$$\Gamma(t,x,y) = (2\pi t)^{-d/2} e^{-\|x-y\|^2/2t}.$$

Let D be a bounded domain with C^3 boundary. Let $X = \{X_t, t \ge 0\}$ be the standard reflecting Brownian motion on D and $L = \{L(t), t \ge 0\}$ be its boundary local time. We refer to [3] for a discussion of reflecting Brownian motion and the boundary local time.

The condition that the restriction of q on D lies in the class

K_d can be characterized in terms of reflecting Brownian motion.

THEOREM 2. *Let* q *be measurable on* $\mathbb{R}^d$, *then* $qI_D \in K_d$ *if and only if*

$$\lim_{t\to 0} \sup_{x\in\bar{D}} E^x \left[\int_0^t |q|(X_s)\, ds \right] = 0 \tag{4}$$

We delay the proof of Theorem 2 to the end of the paper.

Let $qI_D \in K_d$ and put

$$e_q(t) = \exp\left[\int_0^t q(X_s)\, ds \right].$$

This is finite a.s. by (4). Now we define

$$G_q(x) = E^x\left[\int_0^\infty e_q(s)\, dL(s) \right]. \tag{5}$$

G_q will be called the gauge function for the Neumann problem.

Define the semigroup $\{R_t^{(q)},\ t \geq 0\}$ as follows:

$$R_t^{(q)} f(x) = E^x[e_q(t) f(X_t)].$$

Observe that this semigroup is not necessarily sub-Markovian. In the following, A_t and C_t denote constants depending on t. They are not necessarily the same at each appearance.

LEMMA 1. *For any fixed* $t > 0$, *there is a constant* A_t *such that*

$$\forall f \in L^1(D):\quad \|R_t^{(q)} f\|_\infty \leq A_t \|f\|_1. \tag{6}$$

PROOF. The proof is the same as that for killed Brownian motion given in [1]. The conditions used there are also satisfied by the

reflecting Brownian motion. □

Lemma 1 is used in the next lemma to obtain an inequality in the opposite direction.

LEMMA 2. *Let* $f \geq 0$ *be measurable on* $\bar{D}$. *For any fixed* $t > 0$, *there is a constant* A_t *such that*

$$\forall x \in \bar{D}: \quad \int_{\bar{D}} f(y)dy \leq A_t R_t^{(q)} f(x). \tag{7}$$

PROOF. By (6), with $-q$ for q, and the Schwarz inequality,

$$\begin{aligned} E^x[f(X_t)]^2 &= E^x[\, e_{\frac{1}{2}q}(t)\; f^{\frac{1}{2}}(X_t)\; e_{-\frac{1}{2}q}(t)\; f^{\frac{1}{2}}(X_t)]^2 \\ &\leq E^x[e_q(t)\; f(X_t)]\, E^x[e_{-q}(t)\; f(X_t)] \\ &= R_t^{(q)} f(x)\; R_t^{(-q)} f(x) \\ &\leq A_t\; \|f\|_1\; R_t^{(q)} f(x). \end{aligned} \tag{8}$$

On the other hand, for any $t > 0$, there is a positive constant C_t such that

$$\forall (x,y) \in \bar{D} \times \bar{D}: \quad p(t,x,y) \geq C_t,$$

where $p(t,x,y)$ is the transition density function of the reflecting Brownian motion X. Thus,

$$E^x[f(X_t)] \geq C_t \int_{\bar{D}} f(y)dy. \tag{9}$$

By (8) and (9),

$$C_t^2\; \|f\|_1^2 \leq A_t\; \|f\|_1\; R_t^{(q)} f(x). \tag{10}$$

Hence, if $\|f\|_1 < \infty$, then

$$\|f\|_1 \leq A_t \; R_t^{(q)} f(x). \tag{11}$$

In general, we can replace f by $f \wedge n$ in (11) and apply the monotone convergence theorem. The lemma is proved. □

THEOREM 3. *Let* $q \in K_d$ *and* G_q *be the gauge function defined in* (5). *If* $G_q \not\equiv \infty$, *then it is continuous on* $\bar{D}$, *hence bounded on* $\bar{D}$.

PROOF. By the Markov property,

$$\infty \geq G_q(x) = E^x\Big[\int_0^t e_q(s) \; dL(s) \Big] + E^x[e_q(t) G_q(X_t)]. \tag{12}$$

For any fixed $t > 0$,

$$E^x\Big[\int_0^t e_q(s) \; dL(s)\Big]^2 \leq E^x[e_{|q|}(t) \; L(t)]^2 \leq E^x[e_{2|q|}(t)] \; E^x[L(t)^2]. \tag{13}$$

The first factor in the last member of (13) is bounded by Khas'minskii's lemma (see [1]) and Theorem 2. It is easy to show that

$$\sup_{x\in\bar{D}} E^x[L(t)^2] \leq 2(\sup_{x\in\bar{D}} E^x[L(t)])^2. \tag{14}$$

Since

$$E^x[L(t)] = \int_0^t \int_{\partial D} p(s,x,y) \; d\sigma(y) \; ds \leq \int_0^t \frac{C}{\sqrt{t}} \, ds = 2C\sqrt{t}$$

for a constant C (see [3]), it follows that the second factor in the last member of (13) is also bounded. Hence

$$E^x\Big[\int_0^t e_q(s) \; dL(s)\Big] \text{ is bounded on } \bar{D}. \tag{15}$$

Now suppose $G_q(x_0) < \infty$. By (12),

$$\infty > G_q(x_0) \geq R_t^{(q)} G_q(x_0).$$

By Lemma 2,

$$R_t^{(q)} G_q(x_0) \geq A_t \, \|G_q\|_1.$$

Hence $\|G_q\|_1 < \infty$. By Lemma 1,

$$\|R_t^{(q)} G_q\|_\infty \leq A_t \, \|G_q\|_1.$$

This shows that $R_t^{(q)} G_q$ is bounded. It follows from (12) and (13) that G_q is bounded on $\bar{D}$.

Furthermore, it is known that the semigroup $\{R_t^{(q)}, t > 0\}$ is strong Feller, hence $R_t^{(q)} G_q$ is continuous. Now it follows from (13) and (14) that

$$\lim_{t \to 0} \sup_{x \in \bar{D}} E^x\left[\int_0^t e_q(s)\, dL(s)\right] = 0. \tag{16}$$

Hence by (12), G_q is continuous on $\bar{D}$. The theorem is proved. □

It remains to complete the

PROOF OF THEOREM 2. It was proved in [3] that the transition density function $p(t,x,y)$ of the standard reflecting Brownian motion on D can be written in the form

$$p(t,x,y) = p_0(t,x,y) + p_1(t,x,y), \tag{17}$$

where p_0 and p_1 have the following properties.

(a) There are positive constants c_1, c_2, and a such that

$$(18) \qquad c_2\ \Gamma(t,x,y) \le p_0(t,x,y) \le c_1\ \Gamma(at,x,y).$$

(b) $p_1(t,x,y)$ has the form

$$p_1(t,x,y) = \int_0^t \int_D p_0(t-u,x,z)\ f(u,z,y)\ dz\ du,$$

with

$$\sup_{y\in R^d} \int_D |f(t,x,y)|\ dx \le \frac{C}{\sqrt{t}},$$

where C is a constant.

Now let

$$M_0(t) = \sup_{z\in R^d} \int_0^t \int_D p_0(s,y,z)\ |q|(y)\ dy\ ds.$$

We have

$$(19) \quad \int_0^t \int_D |p_1(s,y,x)|\ |q|(y)\ dy\ ds$$

$$\le \int_0^t \int_D \int_0^s \int_D p_0(s-u,y,z)\ |f(u,z,x)|\ |q|(y)\ dz\ du\ dy\ ds$$

$$\le M_0(t) \sup_{x\in \bar{D}} \int_0^t \int_D |f(u,z,x)|\ dz\ ds$$

$$\le 2C\sqrt{t}\ M_0(t).$$

By the symmetry of $p(t,x,y)$ in (x,y), we have

$$E^x\left[\int_0^t |q|(X_s)\ ds\right] = \int_0^t \int_D p(s,y,x)\ |q|(y)\ dy\ ds$$

$$= \int_0^t \int_D p_0(s,y,x)\ |q|(y)\ dy\ ds$$

$$+ \int_0^t \int_D p_1(s,y,x)\ |q|(y)\ dy\ ds.$$

The absolute value of the last term is not greater than $M_0(t)/2$ if

$t \leq 1/8C^2$. Hence for $t \leq 1/8C^2$, we have

$$\frac{1}{2}M_0(t) \leq \sup_{x\in D} E^x\left[\int_0^t |q|(X_s)\, ds\right] \leq \frac{3}{2}M_0(t). \tag{20}$$

On the other hand, as recalled before, $qI_D \in K_d$ if and only if (3) holds. By (18), the latter condition is equivalent to the condition $\lim_{t\to 0} M_0(t) = 0$. Hence by (20), $qI_D \in K_d$ is equivalent to (4). The theorem is proved.

We refer to [3] for further properties of the gauge function G_q as well as its application to the Neumann problem.

References

1. M. AIZENMAN and B. SIMON. Brownian Motion and Harnack's inequality for Schrödinger Operators. *Comm. Pure Appl. Math.* *35*(1982), 209-271.

2. K.L. CHUNG and K.M. RAO. Feynman-Kac Functional and the Schrödinger Equation. *Seminar on Stochastic Process 1981*, pp. 1-29. Birkhäuser, Boston, 1981.

3. PEI HSU. Reflecting Brownian Motion, Boundary Local Time and the Neumann Problem. Doctoral Dissertation, Stanford University, June 1984.

4. Z.X. ZHAO. Conditional Gauge with Unbounded Potentials. Z. *Wahrscheinlichkeitstheorie verw. Geb.* *35*(1983), 13-18.

K.L. CHUNG
Department of Mathematics
Stanford University
Stanford, CA 94305

Pei HSU
Courant Institute of Mathematics
New York University
251 Mercer Street
New York, NY 10012

Seminar on Stochastic Processes, 1985
Birkhäuser, Boston, 1986
With kind permission of Springer Science and Business Media.

DOUBLY-FELLER PROCESS WITH MULTIPLICATIVE FUNCTIONAL

by

K. L Chung*

Despite the common use of the term "Feller property", there are variations in its definition. In the early literature on Markov processes, there are discussions of this and related properties, often under sets of bewildering assumptions. The coast should now be clear, but certain neat formulations may have been overlooked. In §1 of this note, some old results are reviewed in more general forms and an apparently new one is derived. In §2, the results are extended to include a multiplicative functional, of which the prime example is that of Feynman-Kac, properly generalized.

1. Doubly-Feller Process

Let E be a locally compact separable metric space, $E_\Delta = E \cup \{\Delta\}$ its one-point compactification, $\mathscr{E}$ its Borel

*This work was supported in part by NSF grant DMS83-01072 at Stanford University.

tribe. Any function f defined on E is extended to E_Δ by setting $f(\Delta) = 0$.

$b\mathscr{E}$ = the class of bounded $\mathscr{E}$-measurable functions on E;

bC = the class of bounded continuous functions on E;

C_0 = the class of continuous functions on E such that its extension to E_Δ (as specified above) is continuous in E_Δ.

For $f \in b\mathscr{E}$, we write $\|f\| = \sup\{|f(x)| : x \in E\}$,

Let $\{P_t, t > 0\}$ be a probability transition semigroup with state space E_Δ; P_0 being the identity. It is said to have the Feller property if the following two conditions are satisfied:

(i) for each $t > 0$ and $f \in C_0$, we have $P_t f \in C_0$;

(ii) for each $f \in C_0$, $\lim\limits_{t \downarrow 0} P_t f(x) = f(x)$.

It is known that (i) and (ii) together imply that for each $f \in C_0$,

(ii') $$\lim_{t \downarrow 0} \|P_t f - f\| = 0.$$

The semigroup (P_t) is said to have the strong Feller property if

(iii) for each $t > 0$ and $f \in b\mathscr{E}$, $P_t f \in bC$.

It is well known ([1], Chapter 2) that if (P_t) has the

Feller property, then a Hunt process $X = \{X_t, t \geq 0\}$ can be constructed with (P_t) as its transition semigroup. Thus the paths are right continuous and quasi left continuous. Such a process will be called a Feller process. Let $\{\mathcal{F}_t^{\sim}, t \geq 0\}$ be the augmented natural filtration. For each $B \in \mathcal{E}$, define

$$T_B = \inf\{t > 0 : X_t \in B\};$$

$$\tau_B = T_{E-B}.$$

For a Hunt process, it is known that for each $t > 0$, $\{T_B < t\} \in \mathcal{F}_t^{\sim}$, and the function $x \to P^x\{T_B < t\}$ is universally measurable, denoted by $\mathcal{E}^{\sim}$. A full discussion of these questions of measurability is given in [1], §2.3. We begin with a useful consequence.

LEMMA 1. If (P_t) has the strong Feller property, then for each $t > 0$ and $f \in b\mathcal{E}^{\sim}$ (bounded and in $\mathcal{E}^{\sim}$), we have $P_t f \in bC$.

PROOF. (see [3], Annexe 5): Let $\{x_n\}$ be a dense set in E, and

$$\lambda = \sum_{n=1}^{\infty} 2^{-n} \varepsilon_{x_n}$$

where ε_x is the point mass at x. Given $f \in b\mathcal{E}^{\sim}$, there exist f_1 and f_2 in $\mathcal{E}$ such that $f_1 \leq f \leq f_2$ and $(\lambda P_t)(f_1) = (\lambda P_t)(f_2)$. Hence $P_t f_1 \leq P_t f \leq P_t f_2$ and $\lambda(P_t f_2 - P_t f_1) = 0$. Since $P_t f_1$ and $P_t f_2$ are continuous and the measure λ charges every nonempty open set, it follows

that $P_t f_1 = P_t f_2$, and consequently $P_t f = P_t f_1 \in bC$.

The following result is proved in [2] (or see [1], p. 73, Exercise 2).

LEMMA 2. *Let* X *be a Feller process. For each nonempty open set* B *and compact subset* K *of* B, *we have*

$$\lim_{t\to 0}\sup_{x\in K} P^x\{\tau_B \le t\} = 0. \tag{1}$$

REMARK. A condition like (1) can be found in [4] under assumption of continuous paths, but it was not deduced from the "Feller property" which was defined differently there.

The next result is given for the Brownian motion and an open set B in at least four textbooks, including [1]. Since the general case is not stated, we repeat its known proof here to illustrate the measurability question.

LEMMA 3. *Let* X *be a Hunt process whose transition semigroup* (P_t) *has the strong Feller property. Then for each* $B \in \mathcal{E}$, *both functions below are upper semi-continuous:*

$$x \to P^x\{t < T_B\}, \qquad x \to P^x\{t \le T_B\}. \tag{2}$$

PROOF. We have the fundamental relation:

$$T_B = \lim_{s\downarrow 0}\downarrow (s + T_B \circ \theta_s)$$

where θ_s is the shift operator. It follows that

$$(3)\quad P^x\{t < T_B\} = \lim_{s\downarrow 0} \downarrow P^x\{t - s < T_B \circ \theta_s\}$$

$$= \lim_{s\downarrow 0} \downarrow P^x\{P^{X_s}[t - s < T_B]\} = \lim_{s\downarrow 0} P_s\phi_s(x)$$

where

$$\phi_s(x) = P^x[t - s < T_B].$$

Hence $\phi_s \in b\mathcal{E}^{\sim}$ because X is a Hunt Process, and $P_s\phi_s \in bC$ by Lemma 1. It follows from (3) that the first function in (2) is upper semi-continuous. So is the second by the same proof, changing "<" into "≤" in the obvious places.□

A doubly-Feller process is a Feller process whose transition semigroup has also the strong Feller property. The most famous example is the Brownian motion in R^d. Most diffusion processes are doubly-Feller processes. But we are not assuming continuous paths here.

Let B be a nonempty open subset of E, $B \neq E$; and let $B \cup \{\Delta_B\}$ be its one-point compactification, where $\Delta_B \neq \Delta$. Define

$$(4)\qquad X_t^B = \begin{cases} X_t & \text{on } \{t < \tau_B\}; \\ \Delta_B & \text{on } \{t \geq \tau_B\}; \end{cases}$$

The process $X^B = \{X_t^B,\ t \geq 0\}$ is called "the process X killed outside B." Its state space is $B \cup \{\Delta_B\}$, and its transition semigroup $\{P_t^B,\ t \geq 0\}$ is given by

$$P_t^B(x,A) = E^x\{t < \tau_B; X_t \in A\} \text{ if } x \in B,\ A \in B \cap \mathscr{E};$$

(5)

$$P_t^B(x,\{\Delta_B\}) = 1 - P_t^B(x,E) \qquad \text{if } x \in B;\ P_t^B(\Delta_B,\{\Delta_B\}) = 1.$$

If X is a Hunt process, it can be verified that X^B is also a Hunt process. There is no difficulty with the right continuity and quasi left continuity of paths, while the strong Markov property is shown exactly as in Theorem 2 of §4.5 of [1], which treats the special case where X is the Brownian motion.

We denote by $b\mathscr{E}(B)$, $b\mathscr{E}^\sim(B)$, $bC(B)$ the indicated classes of functions restricted to B. Let B^* be the boundary of B in E_Δ. When f is defined only on B, we extend it to E_Δ by setting it to be zero outside B. If this extension is continuous in E_Δ we say that f belongs to $C_0(B)$. The open set B is said to be regular if for each $z \in B^* \cap E$, we have $P^z\{\tau_B = 0\} = 1$. This is the definition used in the Dirichlet problem for B, but note that "regularity at Δ" is not defined when $\Delta \in B^*$.

The following theorem is the main result of this section. It is known when X is the Brownian motion in R^d. In the general form given here it is apparently new, although analogous results may be found in the literature.

THEOREM. Let X be a doubly-Feller process with the state space E_Δ. Let B be a nonempty proper subset of E which is open and regular, and define the process X^B by (4). Then X^B is also a doubly-Feller process.

PROOF: As remarked above, X^B is a Hunt process with the transition semigroup (P_t^B) given by (5). We prove first that the latter has the strong Feller property. Let $t > 0$, $f \in b\mathscr{E}(B)$. Then by (5),

$$(6) \qquad P_t^B f(x) = E^x\{t < \tau_B;\ f(X_t)\}, \qquad x \in B.$$

For $0 < s < t$, we have by the Markov property:

$$(7) \qquad P_t^B f(x) = E^x\{s < \tau_B;\ \phi_s(X_s)\},$$

where

$$\phi_s(x) = E^x\{t - s < \tau_B;\ f(X_{t-s})\}, \qquad x \in E.$$

Then $\phi_s \in b\mathscr{E}^{\sim}$ and so by Lemma 1, $P_s\phi_s \in bC$. It follows from (7) that

$$(8) \quad \left|P_t^B f(x) - P_s\phi_s(s)\right| \le P^x\{\tau_B \le s\}\|\phi_s\|, \qquad x \in B.$$

Since $\|\phi_s\| \le \|f\|$ for all s, the right member of (8) converges to zero uniformly for x in every compact subset of B by Lemma 2. Hence $P_t^B f \in bC(B)$. This proves the strong Feller property of (P_t^B).

REMARK. If we change "$t < \tau_B$" in the definition of the function in (6) into "$t \le \tau_B$", the resulting function is also continuous in B by the same argument, provided that $f \in b\mathscr{E}$.

Next we prove that if $t > 0$ and $f \in C_0(B)$, then

$P_t^B f \in C_0(B)$. Since B^* may contain Δ, let us consider first this case. We have then as $x \in B$, $x \to \Delta$:

$$P_t^B|f|(x) \leq P_t|f|(x) \to 0$$

since $|f| \in C_0$, where f is the extension of f to E_Δ, specified above. On the other hand, if $z \in B^* \cap E$, then we have by Lemma 3 and the regularity assumption

$$\text{(9)} \qquad \overline{\lim_{x \to z}}\, P^x\{t < \tau_B\} \leq P^z\{t < \tau_B\} = 0;$$

and consequently as $x \in B$, $x \to z$,

$$|P_t^B f(x)| \leq P^x\{t < \tau_B\}\|f\| \to 0.$$

We have therefore proved that (P_t^B) satisfies condition (i) of the Feller property.

Finally, let $f \in C_0(B)$ and extend it to E_Δ as before. We have then obviously

$$\text{(9)} \qquad |P_t^B f(x) - P_t f(x)| \leq P^x\{\tau_B \leq t\}\|f\|.$$

For each $x \in B$, $P^x\{\tau_B = 0\} = 0$ since B is open and the paths are right continuous. Hence it follows from (9) that

$$\lim_{t \downarrow 0} P_t^B f(x) = \lim_{t \downarrow 0} P_t f(x) = f(x), \qquad x \in B$$

by the Feller property of (P_t). Thus (P_t^B) satisfies condition (ii) of the Feller property, and we have concluded the proof that X^B is a doubly-Feller process.

COROLLARY. <u>For a doubly-Feller process, and a nonempty open set</u> B, <u>the two functions in</u> (2) <u>are both continuous in</u> B, <u>indeed in</u> E <u>if</u> B <u>is regular</u>.

PROOF: The first function is just $P_t^B 1_B$. The assertion for the second function is a consequence of the Remark in the preceding proof.

Even in the case of a Brownian motion, the two results in the Corollary seem to have escaped notice in the literature.

2. Multiplicative Functional

Let $M = \{M_t, t > 0\}$ be a multiplicative functional associated with X. Namely, for each $x \in E$, P^x-almost surely: $M_0 \equiv 1$, $0 < M_t < \infty$, $M_t \in \tilde{\mathscr{F}}_t$ for $t > 0$; and

$$(10) \quad M_{t+s} = M_s \cdot (M_t \circ \theta_s), \qquad \text{for all } t > 0,\ s > 0.$$

We now impose a set of special conditions on M as follows.

(a) For some $t > 0$:

$$\sup_{x \in E} \sup_{0 \leq s \leq t} E^x\{M_t\} < \infty.$$

It follows from this, condition (10) and the Markov property that for $x \in E$ and $s \in [0,t]$:

$$E^x\{M_{t+s}\} = E^x\{M_s E^{X_s}[M_t]\} \leq \sup_{x \in E} \sup_{0 \leq s \leq t} E^x\{M_s\}^2.$$

Hence by induction (a) is in fact true for all (finite) t. We shall denote the bound by A_t below.

(b) For each $t > 0$, there exists a number $\alpha > 1$ (which may depend on t) such that

$$\sup_{x \in E} E^x\{M_t^\alpha\} < \infty.$$

(c) For each compact subset K of E, we have

$$\lim_{t \to 0} \sup_{x \in K} E^x\{|M_t - 1|\} = 0.$$

These conditions are the simplest to yield Theorem 2 below. They are inspired by our prime example which we take up first.

EXAMPLE. Let $q \in \mathscr{E}$, and suppose that

$$\lim_{t \to 0} \sup_{x \in E} E^x\{\int_0^t |q|(X_s)ds\} = 0. \tag{11}$$

Put

$$e_q(t) = \exp\{\int_0^t q(X_s)ds\}.$$

Then $\{e_q(t), t > 0\}$ is a multiplicative functional satisfying all the conditions above. It is known as the Feynman-Kac functional when X is the Brownian motion. The condition (11) is due to Aizenman and Simon.

It follows from (11), (10) and the Markov property that the sup in (11) is in fact finite for all $t > 0$. The argument is similar to that given under (a), changing

multiplication to addition. Hence for each x, P^x-almost surely, we have $0 < e_q(t)\ \infty$, and $e_q(t) \in \mathscr{F}_t^{\sim}$. It is easy to verify (10) for $M_t = e_q(t)$. (Indeed $\{\int_0^t q(X_s)ds,\ t > 0\}$ is an additive functional.) To verify the conditions (a), (b) and (c), let us cite the following lemma due to Khas'minskii [5].

LEMMA 4. <u>If for a fixed t the</u> sup <u>in</u> (11) <u>is less than</u> $\varepsilon < 1$, <u>then</u>

$$\sup_{x\in E} E^x\{e_{|q|}(t)\} < \frac{1}{1-\varepsilon}. \tag{12}$$

The proof is by using Taylor's series for the exponential function, then estimating each $E^x\{(\int_0^t |q|(X_s)ds)^n\}$ by converting the n^{th} power into an n - tuple integral, and using the Markov property.

Therefore under (11), condition (a) is satisfied by $M_t = e_{|q|}(t)$ for a sufficiently small $t > 0$ since $e_{|q|}(t)$, is increasing with t. Since $\sup\{e_q(s):0 < s < t\} < e_{|q|}(t)$ this implies condition (a) for $M_t = e_q(t)$. Next, for any real constant α, we have

$$\{(e_q(t))^\alpha = e_{\alpha q}(t) < e_{|\alpha q|}(t). \tag{13}$$

Now (11) remains true when q is replaced by αq; hence condition (b) is satisfied by $M_t = e_q(t)$, indeed for any $\alpha > 1$ independent of t. Finally, let $\varepsilon > 0$ and choose $t > 0$ so that the sup in (12) is less that $1 + \varepsilon$. For this fixed t let $\Lambda = \{e_q(t) > 1\}$. Then for all $x \in E$ we have

$$(14)\qquad E^x\{|e_q(t) - 1|\} = E^x\{e_q(t) - 1;\ \Lambda\} + E^x\{1 - e_q(t);\ \Lambda^c\}.$$

The first term on the right side is bounded by $E^x\{e_{|q|}(t) - 1;\ \Lambda\}$. The second term is bounded by

$$E^x\{1 - e_{-|q|}(t);\ \Lambda^c\} \leq E^x\{e_{|q|}(t) - 1;\ \Lambda^c\}$$

because $e_{|q|}(t) \geq 1$. Adding up we see the left member in (14) is bounded by $E^x\{e_{|q|}(t)\} - 1 < \varepsilon$. This establishes a stronger form of the condition (c) for $M_t = e_q(t)$.

Now we proceed to extend the results in §1 to a process with multiplicative functional. We begin by defining $\{Q_t,\ t \geq 0\}$ as follows: for $f \in b\mathscr{E}$:

$$(15)\qquad Q_t f(x) = E^x\{M_t f(X_t)\}, \qquad x \in E.$$

By means of (10), we can verify that $\{Q_t\}$ forms a semigroup, not necessarily submarkovian. But we have for each $t \geq 0$,

$$(16)\qquad \|Q_t f\| \leq \sup_{x\in E} E^x\{M_t\}\|f\| \leq A_t\|f\|$$

by (a), so that each Q_t maps $b\mathscr{E}$ into $b\mathscr{E}$. The definitions of both Feller properties can be extended to $\{Q_t\}$ without change. We shall keep the notation (P_t) for the semigroup of the process X.

THEOREM 2. <u>If</u> (P_t) <u>has both Feller properties, then so does</u> (Q_t) <u>provided that</u> (M_t) <u>satisfies the conditions</u> (a), (b) <u>and</u> (c).

PROOF: We have by using (10) together with the Markov property, for $x \in E$ and $f \in b\mathscr{E}$:

$$Q_t f(x) = E^x\{M_s \cdot [M_{t-s} f(X_{t-s})] \circ \theta_s\}$$

$$= E^x\{M_s Q_{t-s} f(X_s)\}.$$

Hence

$$(17) \quad |Q_t f(x) - P_s Q_{t-s} f(x)| = |E^x\{(M_s - 1) Q_{t-s} f(X_s)\}|$$

$$\leq E^x\{|M_s - 1|\} \| Q_{t-s} f \| \leq E^x\{|M_s - 1|\} A_t \| f \|$$

by condition (a) and an estimate like (16). Since $Q_{t-s} f \in b\mathscr{E}^{\sim}$, we have $P_s Q_{t-s} f \in bC$ by Lemma 1. Letting $s \to 0$ in (17), the right member converges to zero uniformly in each compact by condition (c). Hence $Q_t f \in bC$, and we have verified the strong Feller property for (Q_t).

Next, if $f \in C_0$, then so is $|f|^\alpha$ for any $\alpha > 0$. We have by Hölder's inequality applied to (15):

$$(18) \qquad |Q_t f(x)| \leq E^x\{M_t^\alpha\}^{1/\alpha} E^x\{|f|^{\alpha'}(X_t)\}^{1/\alpha'}$$

where $\alpha^{-1} + (\alpha')^{-1} = 1$. By condition (b), the first factor on the right side of (18) is bounded in x; as $x \to \Delta$, the second factor converges to zero because (P_t) has the Feller

property. Hence $Q_t f \in C_0$.

Finally, if $f \in C_0$, then

$$|Q_t f(x) - P_t f(x)| \leq E^x\{|M_t - 1|\}\|f\|$$

and consequently by a weaker form of condition (c), for each $x \in E$:

$$\lim_{t\to 0} Q_t f(x) = \lim_{t\to 0} P_t f(x) = f(x).$$

Thus (Q_t) satisfies the conditions of the Feller property.□

Combining Theorems 1 and 2, we obtain

THEOREM 3. Let X be a doubly-Feller process, B as in Theorem 1, and M as in Theorem 2. Define for $x \in E$, $f \in b\mathscr{E}$.

$$(19) \qquad Q_t^B f(x) = E^x\{t < \tau_B;\ M_t f(X_t)\}.$$

Then $\{Q_t^B,\ t > 0\}$ is a doubly-Feller semigroup.

COROLLARY. If, in addition, B is relatively compact, then we have $Q_t^B f \in C_0(B)$ for each $t > 0$ and $f \in b\mathscr{E}(B)$.

PROOF: In this case $\Delta \notin B^*$. Let $z \in B^*$, then we have by (19):

$$|Q_t^B f(x)| \leq E^x\{t < \tau_B\}^{1/\alpha'} E^x\{M_t^\alpha |f|^\alpha (X_t)\}^{1/\alpha}.$$

Hence as $x \to z$, the above converges to zero by (9) and condition (b). Thus $Q_t^B f \in C_0(B)$.

The following alternative approach to Theorem 3 is illuminating. It is well known that the killing operation is representable by a multiplicative functional as follows:

$$\tilde{M}_t = 1_{\{t < \tau_B\}}, \qquad t \geq 0,$$

provided that the original state space E is replaced by the new state space B, as it should be because X_t lives on B on $\{t < \tau_B\}$. It is easy to verify then all the conditions imposed on $\tilde{M} = \{\tilde{M}_t\}$ at the beginning of this section, starting with the fundamental relation (10) which holds P^x-almost surely for all $x \in B$. Conditions (a) and (b) are trivial while condition (c) reduces precisely to Lemma 2. Now Theorem 3 can be deduced from Theorem 2 by applying it to the double multiplicative functional

$$\tilde{M}_t M_t = 1_{\{t<\tau_B\}} M_t, \qquad t \geq 0,$$

where M_t is as in Theorem 2.

References

[1] K. L. Chung, Lectures from Markov Processes to Brownian Motion, Grundlehren der mathematischen Wissenschaften 249, Springer-Verlag 1982.

[2] K. L. Chung and R. K. Getoor, The condenser problem, Ann. Probab. 5 (1977), 82-86.

78

[3] P. Courrège and P. Priouret, Axiomatique du problème de Dirichlet et processus de Markov, Sém. Brelot-Choquet-Deny (Théorie de potentiel), 8e année, 1963/64.

[4] I. V. Girsanov, Strongly-Feller processes, I. General properties, Theory Probab. Appl. (translated from the Russian) 5 (1960), 5–24.

[5] R. Z. Khas'minskii, On positive solutions of the equation Au + Vu = 0, Theory Probab. Appl. (translated from the Russian) 4 (1959), 309–318.

K. L. Chung
Department of Mathematics
Stanford University
Stanford, California 94305

TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 306, Number 2, April 1988
With permission from the American Mathematical Society.

GENERAL GAUGE THEOREM FOR MULTIPLICATIVE FUNCTIONALS

K. L. CHUNG AND K. M. RAO

ABSTRACT. We generalize our previous work on the gauge theorem and its various consequences and complements, initiated in [8] and somewhat extended by subsequent investigations (see [6]). The generalization here is two-fold. First, instead of the Brownian motion, the underlying process is now a fairly broad class of Markov processes, not necessarily having continuous paths. Second, instead of the Feynman-Kac functional, the exponential of a general class of additive functionals is treated. The case of Schrödinger operator $\Delta/2+\nu$, where ν is a suitable measure, is a simple special case. The most general operator, not necessarily a differential one, which may arise from our potential equations is briefly discussed toward the end of the paper. Concrete instances of applications in this case should be of great interest.

1. Preliminaries. We begin with a Hunt process $(X_t, \mathcal{F}_t, \theta_t; t \geq 0)$ on $(E, \mathcal{E})$ with transition semigroup (P_t); see Chung [2] for terminology and notations.

Let

$\mathcal{E}_+ =$ the class of $\mathcal{E}$-measurable functions which are positive (≥ 0).

$L^\infty =$ the class of bounded $\mathcal{E}$-measurable functions.

$C_b =$ the class of bounded continuous functions.

$C_0 =$ the class of continuous functions vanishing at ∂ where ∂ is the one-point compactification of E. If open $D \subset E$, then $L^\infty(D)$ and $C_b(D)$ will denote respectively the restriction of L^∞ and C_b to D; while $C_0(D)$ the class of continuous functions on D vanishing at $\partial_D =$ the one-point compactification of D.

For any $A \in \mathcal{E}$, let

$$T_A = \inf\{t > 0 \colon X_t \in A\}, \qquad \tau_A = \inf\{t > 0 \colon X_t \in A^C\},$$
$$P_A f(x) = E^x\{f(X(T_A)); T_A < \infty\}.$$

We assume that (P_t) has the strong Feller property, i.e., for each $t > 0$ and $f \in L^\infty$, $P_t f \in C_b$. The following lemma is elementary, but will be proved.

Let U^λ, $\lambda \geq 0$, denote the λ-potential of (P_t),

$$U^\lambda f(x) = \int_0^\infty e^{-\lambda t} P_t f \, dt; \qquad 0 \leq f \in L^\infty. \tag{1}$$

LEMMA 1. *If (P_t) has the strong Feller property, then for each $\lambda > 0$, and $f \in L^\infty$, we have*

$$U^\lambda f \in C_b. \tag{2}$$

Received by the editors March 9, 1987.

1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 60J45, 60J56; Secondary 60J40, 35J10.

The first author's research was supported in part by AFOSR grant 85-0330 and NSF grant DMS 83-01072.

Suppose there is a Radon measure m on $\mathcal{E}$ such that $U^\lambda(x,\cdot) \ll m(\cdot)$. Let K be a compact subset of E. Then for any $\lambda > 0$ the set of measures $\{U^\lambda(x,\cdot), x \in K\}$ is equi-absolutely continuous with respect to m, namely, for any $\varepsilon > 0$ there exists $\delta > 0$ such that for all $A \in \mathcal{E}$ with $m(A) < \delta$ we have

$$\sup_{x\in K} U^\lambda(x,A) < \varepsilon. \tag{3}$$

PROOF. First, (2) is trivial by the dominated convergence theorem applied to (1). It follows from (1) and Dini's theorem that if $A_n \in \mathcal{E}$, $A_n \downarrow$ and $m(A_n) \downarrow 0$, then

$$\sup_{x\in K} U^\lambda(x,A_n) \downarrow 0. \tag{4}$$

Now suppose (3) is false, then there exists $\varepsilon > 0$, $x_n \in K$, $m(A_n) < 2^{-n}$ such that $U^\lambda(x_n, A_n) \geq \varepsilon$; a fortiori,

$$U^\lambda\left(x_n, \bigcup_{k=n}^\infty A_k\right) \geq \varepsilon.$$

Let $x_n \to x \in K$, then by continuity,

$$U^\lambda\left(x, \bigcup_{k=n}^\infty A_k\right) \geq \varepsilon$$

which is a contradiction to (4).

From now on the measure m in Lemma 1 will be called a reference measure. The assumption that there exists $\lambda > 0$ such that for every x, $U^\lambda(x,\cdot)$ is absolutely continuous with respect to m will hold throughout this paper. Using the resolvent equation it is seen that if this is true for some $\lambda \geq 0$, then it is true for all $\lambda \geq 0$. But we need not assume the finiteness of U^0. Lemma 1 is used to prove the next result which is essential.

LEMMA 2. *Let D be a relatively compact subset of E. For each $\varepsilon > 0$ there exists $\delta > 0$ such that for all $B \in \mathcal{E}$, $B \subset D$ with $m(B) < \delta$*

$$\sup_{x\in\overline{B}} E^x\{\tau_B\} < \varepsilon.$$

PROOF. We have for each x, and $0 < \varepsilon < 1/2$,

$$\begin{aligned}
\varepsilon \sup_{x\in\overline{D}} P^x\{\tau_B > \varepsilon\} &\leq \sup_{x\in\overline{D}} e \int_0^\varepsilon e^{-t} P^x\{\tau_B > t\}\, dt \\
&\leq \sup_{x\in\overline{D}} e \int_0^\infty e^{-t} P^x\{X_t \in B\}\, dt \\
&= e \sup_{x\in\overline{D}} U^1(x,B).
\end{aligned}$$

By Lemma 1 the last supremum is less than ε^2/e provided that $m(B)$ is small enough. Then by the Markov property

$$\sup_x P^x\{\tau_B > n\varepsilon\} \leq \varepsilon^n$$

and hence for all x

$$E^x\{\varepsilon^{-1}\tau_B\} \le \sum_0^\infty P^x\{\varepsilon^{-1}\tau_B > n\} \le \sum_0^\infty \varepsilon^n = (1-\varepsilon)^{-1},$$

i.e. $E^x[\tau_B] \le \varepsilon(1-\varepsilon)^{-1} < 2\varepsilon$.

COROLLARY. *For $t > 0$, we have*

$$\sup_x P^x\{\tau_B > t\} < \varepsilon/t. \tag{5}$$

A function $f \in \mathcal{E}_+$ is excessive iff $f \ge P_t f$ for $t \ge 0$ and $f = \lim_{t\downarrow 0} P_t f$. An excessive function is called a potential iff $P_{K^c} f$ converges to zero almost everywhere as the compact set K increases to E. Let V be a potential which moreover belongs to C_0. Then it is known [**1**, Chapter 4, Theorems 3.8 and 3.13] that V has a representation as follows:

$$V(x) = E^x\left\{\int_0^\infty dA_t\right\} = E^x\{A_\infty\} \tag{6}$$

where $\{A_t, t \ge 0\}$ is a unique increasing continuous additive functional with $A_0 = 0$. We shall also use the notation

$$U_A f(x) = E^x\left\{\int_0^\infty f(X_t)\, dA_t\right\} \tag{7}$$

for suitable f; so that $V = U_A 1$.

It follows from the definition of additive functional that for $t \ge 0$ and $s \ge 0$

$$A_{t+s} = A_t + A_s \circ \theta_t. \tag{8}$$

Here for each $t \ge 0$:

$$E^x\{A_t\} = V(x) - P_t V(x) \tag{9}$$

is continuous in x because $P_t V$ is by the strong Feller assumption. Since $A_t \downarrow 0$ as $t \downarrow 0$, it follows by Dini's theorem and the assumption $V \in C_0$ that

$$\lim_{t\downarrow 0} E^x\{A_t\} = 0 \tag{10}$$

uniformly in $x \in E$.

2. Define the functional

$$e(t) = e_A(t) = e^{A(t)}. \tag{11}$$

Thus $\{e(t), t \ge 0\}$ is a multiplicative functional. Its first basic property is given by Lemma 3.

LEMMA 3. *There exist constants $C > 1$ and $b > 0$ such that*

$$\sup_{x\in E} E^x\{e(t)\} \le Ce^{bt}. \tag{12}$$

PROOF. By (10), there exists $t_0 > 0$ such that

$$\sup_{x\in E} E^x\{A_{t_0}\} < \theta < 1.$$

Hence for all x, $0 \le t \le t_0$, and $n \ge 2$,

$$\begin{aligned}\frac{1}{n}E^x\{(A_t)^n\} &= E^x\left\{\int_0^t (A_t - A_s)d(A_s)^{n-1}\right\}\\ &= E^x\left\{\int_0^t A_{t-s}(\theta_s)d(A_s)^{n-1}\right\}\\ &= E^x\left\{\int_0^t E^{X_s}[A_{t-s}]d(A_s)^{n-1}\right\}\\ &\le \theta E^x\{(A_t)^{n-1}\}.\end{aligned}$$

Therefore, by induction we have for all x:

$$E^x\{A_t^n\} \le n!\theta^n,$$

$$E^x\{e^{A_t}\} \le \sum_{n=0}^{\infty}\theta^n = \frac{1}{1-\theta}. \tag{13}$$

Since by (8),

$$e^{A_{t+s}} = e^{A_t}(e^{A_s}(\theta_t)),$$

we obtain for $(n-1)t_0 < t \le nt_0$,

$$E^x\{e^{A_t}\} \le 1/(1-\theta)^n.$$

This is equivalent to the assertion in (12).

COROLLARY. *For any $n > 0$,*

$$\lim_{t\to 0}\sup_{x\in E} E^x\{|e(t)-1|^n\} = 0.$$

PROOF OF THE COROLLARY. Observe that for $\alpha > 0$, $e^\alpha - 1 \le \alpha e^\alpha$. Then

$$\begin{aligned}E^x[|e(t)-1|^n] &\le E^x(e(t)^n - 1) = E^x\{e^{nA_t} - 1\} \le E^x\{(nA_t)e^{nA_t}\}\\ &\le nE^x\{A_t^2\}^{1/2}(E^x\{e^{2nA_t}\})^{1/2}.\end{aligned}$$

Observe that $\sup_x E^x\{A_t^2\} \le \sup_x 2E^x\{A_t\}^2 \to 0$ as $t \to 0$ by (10) and

$$\sup_x \sup_{0<t<1} E^x\{e^{2nA_t}\} < \infty$$

by (12) applied to $2nA$.

Lemma 2 will now be sharpened. Note first of all that if $m(B)$ is small enough, then $\tau_B < \infty$ a.s. We shall assume this below.

LEMMA 4. *Under the conditions of Lemma 2, there exists $\delta > 0$ such that if $m(B) < \delta$, then*

$$\sup_{x\in\overline{B}} E^x\{e(\tau_B)\} < \infty. \tag{14}$$

PROOF. If $P^x\{\tau_B = 0\} = 1$, then $E^x\{e(\tau_B)\} = 1$. Otherwise $P^x\{\tau_B = 0\} = 0$ and we have

$$\begin{aligned} E^x\{e(\tau_B)\} &= \sum_{n=0}^{\infty} E^x\{n < \tau_B; e(n)E^{X_n}\{0 < \tau_B \le 1; e(\tau_B)\}\} \\ &\le \sum_{n=0}^{\infty} E^x\{n < \tau_B; e(n)E^{X_n}\{e(1)\}\} \\ &\le C_1 \sum_{n=0}^{\infty} E^x\{n < \tau_B; e(n)\} \end{aligned}$$

where

$$C_1 = \sup E^x\{e(1)\} < \infty$$

by Lemma 3. The sum above is bounded by

$$\sum_{n=0}^{\infty} P^x\{n < \tau_B\}^{1/2} E^x\{e^{2A(n)}\}^{1/2}.$$

Since $\{2A_t\}$ is the additive functional associated with $2V$ which satisfies the same conditions as V, Lemma 3 applied to $\{2A_t\}$ yields

$$\sup_{x \in E} E^x\{e^{2A(n)}\} \le C_2 e^{nb_2},$$

with some $C_2 > 0$ and $b_2 > 0$. Hence the sum in question is bounded by

$$\sum_{n=0}^{\infty} P^x\{n < \tau_B\}^{1/2} C_2^{1/2} e^{nb_2/2}. \tag{15}$$

Now by the Corollary of Lemma 2, there exists $\delta > 0$ such that if $B \subset D$ and $m(B) < \delta$ then for all x

$$P^x\{1 < \tau_B\} \le e^{-2b_2}.$$

It follows by the Markov property that

$$P^x\{n < \tau_B\} \le \exp(-2nb_2).$$

Using this in (15) we see that the series converges.

The proof given above may be easily improved to show that the quantity in (14) converges to 1 as $\delta \downarrow 0$.

3. The gauge theorem. We generalize the assumptions described in §1 to a function $V = V^{(1)} - V^{(2)}$ where each $V^{(i)}$, $i = 1, 2$, satisfies the conditions for V in §1. Let $A^{(i)}$, $i = 1, 2$, be the continuous increasing additive functional generating $V^{(i)}$ as in (6). Thus $A = A^{(1)} - A^{(2)}$ is a continuous additive functional generating V as in (6), where $\{A_t, t \ge 0\}$ is of bounded variation in $[0, \infty)$ in t. Put

$$A^* = A^{(1)} + A^{(2)}. \tag{16}$$

As before we write $e(t) = e_A(t) = e^{A(t)}$.

Let D be a nonempty open subset of E. We define the function

$$g(x) = E^x\{e_A(\tau_D)\}, \qquad x \in D. \tag{17}$$

Note that we are not assuming $\tau_D < \infty$; on $\{\tau_D = \infty\}$, $|A_\infty| \le A^*_\infty$ is almost surely finite by (6). As shown above $g \in \mathcal{E}_+$. We call g the *gauge* for (D, A). This term was introduced in [8] in the special case when X is a Brownian motion and

$$A(t) = \int_0^t q(X_s)\,ds \quad \text{for } q \in L^\infty(D).$$

This case where q is in an appropriate class of functions will be referred to later as the old case. Since then several generalizations have been given, for the latest see [**6**]. The gauge theorem will be proved in the present framework in two stages. In the first stage, our previous assumptions suffice but the conclusion is weaker than in previously known special cases. Another assumption will be needed to bring it to full fruition.

We begin with the process X^D obtained by killing X outside D. Namely, let $D \cup \{\partial\}$ be the one-point compactification of D, and define

$$X_t^D = \begin{cases} X_t & \text{on } \{t < \tau_D\}, \\ \partial & \text{on } \{t \ge \tau_D\}. \end{cases}$$

The transition semigroup of $\{X_t^D, t \ge 0\}$ will be denoted by $\{P_t^D\}$, so that

$$P_t^D f(x) = E^x\{t < \tau_D; f(X_t)\}$$

for $f \in \mathcal{E}_+(D)$ with $f(\partial) = 0$. We write also for $\lambda \ge 0$,

$$U^{D,\lambda} f(x) = \int_0^\infty e^{-\lambda t} P_t^D f(x)\,dx = E^x\left\{\int_0^\infty e^{-\lambda t} f(X_t^D)\,dt\right\}.$$

For $\lambda > 0$, this is just the Green operator for D. We shall call X^D the "trace of X on D". It is a Hunt process; notations such as "excessive", "finely continuous" for the trace will be prefixed by the letter D. We begin with an essential lemma.

LEMMA 5. *The gauge function g is D-finely continuous, and $g \in \mathcal{E}$.*

PROOF. It is known [**1**, Chapter 2] that excessive functions are finely continuous, so is a difference of excessive functions. Now

$$e(\tau_D) - 1 = \int_0^{\tau_D} \exp(A(\tau_D) - A_t)\,dA_t = \int_0^{\tau_D} \exp(A(\tau_D)(\theta_t))\,dA_t$$

so that if $A = A^{(1)} - A^{(2)}$,

$$E^\cdot\{e(\tau_D)\} - 1 + E^\cdot\left\{\int_0^{\tau_D} e(\tau_D)(\theta_t)\,dA_t^{(2)}\right\} = E^\cdot\left\{\int_0^{\tau_D} e(\tau_D)(\theta_t)\,dA_t^{(1)}\right\},$$

i.e.,

$$g(x) - 1 + E^\cdot\left\{\int_0^{\tau_D} g(X_t) dA_t^{(2)}\right\} = E^\cdot\left\{\int_0^{\tau_D} g(X_t)\,dA_t^{(1)}\right\}.$$

The last two functions are D-excessive. It is known that excessiveness implies nearly Borel measurability. Hence under Hypothesis (L), $g \in \mathcal{E}$ [**1**, Chapter 5]. Since $U^{D,\lambda} \le U^\lambda \ll m$ by the assumptions of Lemma 1, (L) holds for X^D.

THEOREM 1. *Suppose D is nonempty open and relatively compact. Then the set $F = \{x \in D\colon g(x) < \infty\}$ is absorbing relative to D, namely, for any $x \in F$, $P^x\{\tau_F < \tau_D\} = 0$. Moreover, g is actually bounded on F.*

PROOF. Let $x \in F$ and K be any compact subset of $D \setminus F$. By the strong Markov property

$$\infty > g(x) \geq E^x\{T_K < \tau_D; e^{A(T_K)} g(X(T_K))\}. \tag{18}$$

Since K is closed, $X(T_K) \in K$; since g is finely continuous by Lemma 5, $g(X(T_K)) = \infty$ on $\{T_K < \tau_D\}$. On the other hand, almost surely,

$$\exp(A(T_K)) = \exp(A^{(1)}(T_K) - A^{(2)}(T_K))$$

is strictly > 0 because $E^x\{A^{(2)}(T_K)\} \leq E^x\{A^{(2)}_\infty\} = V^{(2)}(x) < \infty$ for all x. It follows from (18) that $P^x\{T_K < \tau_D\} = 0$. This being true for all compact subsets K of $D \subset F$ (which belongs to $\mathcal{E}$), we conclude that

$$P^x\{T_{D\setminus F} < \tau_D\} = 0.$$

Thus F is absorbing as asserted.

Next, since m is a Radon measure, $m(F) < \infty$; hence we can choose N so large that $B = D \cap \{N < g < \infty\}$ has measure $m(B) < \delta$, so that Lemma 4 is applicable. Then we have if $x \in B$,

$$g(x) = E^x\{\tau_B = \tau_D; e(\tau_B)\} + E^x\{\tau_B < \tau_D; e(\tau_B) g(X(\tau_B))\}. \tag{19}$$

The first term on the right side of (19) is bounded by M (say) by Lemma 4. P^x-a.s. on $\{\tau_B < \tau_D\}$, $X(\tau_B)$ does not belong to $D \setminus F$ because F is absorbing, hence it must belong to the fine closure of $\{g \leq N\}$. Since g is finely continuous the latter is just $\{g \leq N\}$. Therefore the second term on the right side of (19) is bounded by MN. It follows that on B, g is bounded by $M + MN$; on $F \setminus B$ it is by N. The theorem is proved. (Nothing can be said of the value of g on $(\partial D) \setminus (D^c)^{\text{reg}}$; but see later.)

THEOREM 2. *Suppose in addition to previous assumptions that for some $\lambda > 0$ the set of measures*

$$\{U^{D,\lambda}(x, \cdot), x \in D\} \tag{20}$$

are mutually equivalent. Then either $g \equiv \infty$ in D, or g is bounded in E.

PROOF. We have proved in Theorem 1 that there exists a constant C such that $F = \{g \leq C\}$. Hence F is finely closed and $D \setminus F$ is finely open. If it is not empty, let $x_0 \in D \setminus F$, then $U^{D,\lambda}(x_0, D \setminus F) > 0$. Under the new assumption this implies $U^{D,\lambda}(x, D \setminus F) > 0$ for any $x \in F$, which is impossible by Theorem 1 unless F is empty. Therefore either F or $D \setminus F$ is empty; if the latter, then g is bounded on $D = F$ as just reiterated.

Next, let $x \in \partial D$ and not regular for D^c, then

$$g(x) = \lim_{t \downarrow 0} E^x\{t < \tau_D; e(t) g(X_t)\} \leq \lim_{t \downarrow 0} E^x\{t < \tau_D; e(t)\} C$$

since $g \leq C$ in D. By the Corollary after Lemma 3, this implies $g(x) \leq C$. For x regular for D^c, of course $g(x) = 1$. Hence g is bounded in all of E.

Let us state here the open problem of the extension of the gauge theorem to the case where D is a general open set. The problem is open even in the case where

$$A(t) = \int_0^t q(X_s)\,ds.$$

In order to strengthen the results above we shall need the strong Feller property for (P^D). A recent result by Chung [3] states that if the original (P_t) has both the Feller and strong Feller properties then for each nonempty open set D, (P_t^D) also has both properties provided every point of ∂D is regular for D^c. (We say then D is regular.) In this case of course the process X^D as well as X is a Feller process, a particular case of Hunt process. As in [3] we shall call such a process "doubly Feller". Under this assumption Lemma 5 has a complement.

LEMMA 6. *If (P_t) is doubly Feller, then the gauge function g is lower semicontinuous in E.*

PROOF. Note that since P_t^D is strong Feller for each bounded Borel function f, $E^{\cdot}\{t < \tau_D; e(t)f(X_t)\}$ is also continuous in D. We have

$$\begin{aligned} g(x) &= \lim_{t\downarrow 0} \uparrow E^x\{t < \tau_D; e(\tau_D)\} = \lim_{t\downarrow 0} E^x\{t < \tau_D; e(t)g(X_t)\} \\ &= \lim_{t\downarrow 0} \uparrow \lim_{n\uparrow\infty} \uparrow E^x\{t < \tau_D; e(t)(g(X_t)\wedge n)\} \end{aligned}$$

and the last expectations are continuous in x as noted above.

Returning to Theorem 1, under the new assumption the set $F = \{g < \infty\} = \{g \le C\}$ is closed and $D \setminus F = D \cap \{g = \infty\}$ is open. We can now give another sufficient condition for the gauge theorem.

PROPOSITION 7. *Let the D in Theorem 1 be also connected. Suppose for some $\lambda > 0$,*

$$U^{D,\lambda}(x,dy) = u^{D,\lambda}(x,y)m(dy)$$

and that for all $x \in D$,

$$\underline{\lim}_{y\to x} u^{D,\lambda}(x,y) > 0. \tag{23}$$

Then either $g \equiv \infty$ in D or g is bounded in D.

PROOF. We may suppose that F is not empty and prove that $D \setminus F$ is empty. Otherwise let $x_0 \in \partial(D\setminus F)\cap D$. Then $x \in F$ since F is closed. Since F is absorbing with respect to X^D, we have $U^{D,\lambda}(x_0, D\setminus F) = 0$, hence

$$u^{D,\lambda}(x_0, y) = 0, \quad \text{for } m\text{-a.e. } y \in D\setminus F. \tag{24}$$

Let B be any neighborhood of x_0; then $B \cap (D\setminus F)$ is nonempty open and so $m(B\cap(D\setminus F)) > 0$. Therefore (24) contradicts (23) and the proposition is proved.

The condition (23) seems contrived but actually can be easily verified in concrete cases, such as Brownian motion. More generally let us suppose that $U^0(x,dy) = u(x,y)m(dy)$ and the following conditions are satisfied:

$$\text{For all } x \in E\colon u^0(x,x) = \infty, \tag{25}$$

For each $\varepsilon > 0$, $u(x,y)$ is bounded for all (x,y) satisfying $\rho(x,y) > \varepsilon$, where ρ is a metric for the topology.

Then we have

$$U^{D,0}(x,dy) = u^D(x,y)m(dy),$$
$$u^D(x,y) = u(x,y) - E^x\{u(X(\tau_D),y)\}$$

and it is easy to check from (25) that

$$\lim_{y\to x} u^D(x,y) = +\infty.$$

We are indebted to Liao Ming regarding the conditions (25). Conditions similar to these in (25) were used by Chung and Rao in [8], and in Liao's dissertation (Stanford University, 1984) in a general study of potential theory.

Suppose that (P_t) is doubly Feller. Define a semigroup $\{Q_t, t \geq 0\}$ as

$$Q_t f(x) = E^x\{e(t)f(X_t)\};$$

then define

$$Q_t^D f(x) = E^x\{t < \tau_D; e(t)f(X_t)\}$$

where D is nonempty, open, not necessarily relatively compact. It is proved in Chung [3] that both $\{Q_t\}$ and $\{Q_t^D\}$ are strongly Feller provided $\{e(t)\}$ satisfies the following conditions:

(a) for some $t > 0$,

$$\sup_{x\in E}\sup_{0\leq s\leq t} E^x\{e(s)\} < \infty;$$

(b) for each $t > 0$, there exists $\alpha > 1$ such that

$$\sup_{x\in E} E^x\{e(t)^\alpha\} < \infty;$$

(c) for each compact subset K of E, we have

$$\lim_{t\downarrow 0}\sup_{x\in K} E^x\{|e(t)-1|\} = 0.$$

(d) Q_t is doubly Feller; Q_t^D is doubly Feller provided D is regular.

To verify these conditions, consider $e^*(t) = e_{A^*}(t)$ where A^* is defined in (16). Then

$$\sup_{0\leq s\leq t} e(s) \leq e^*(t).$$

Hence (a) is true by Lemma 3 applied to $e^*(t)$, namely $A^*(t)$ instead of $A(t)$ there. Since $e(t)^\alpha = e_{\alpha A}(t)$, (b) is true by Lemma 3 applied to αA instead of A. (c) is proved in the Corollary after Lemma 3, by replacing A with A^* there.

We can now sharpen Lemma 6, indeed for a more general situation.

THEOREM 3. *Suppose (P_t) is doubly Feller, D is nonempty, open and regular. If the gauge for D is bounded in D, then it is continuous in E.*

PROOF. Using our new notation

$$g(x) = \lim_{t\downarrow 0} Q_t^D g(x)$$

where for each $t > 0$, $Q_t^D g \in C_b$ because $\{Q_t^D\}$ has the strong Feller property and g is assumed to be bounded. Now

$$\begin{aligned}(26)\qquad 0 \leq g(x) - Q_t^D g(x) &= P^x\{t \geq \tau_D; e(\tau_D)\} \leq P^x\{t\geq \tau_D; e^*(t)\}\\ &\leq P^x\{t \geq \tau_D\}^{1/2} E^x\{e^{2A^*(t)}\}^{1/2}.\end{aligned}$$

By Lemma 3 applied to $2A^*$,

$$\sup_{x\in E} E^x\{e^{2A^*(t)}\} \le C(t) < \infty$$

where $C(t)$ is a constant depending on t. Since (P_t) is doubly Feller,

$$\lim_{t\downarrow 0} \sup_{x\in K} P^x\{t \ge \tau_D\} = 0$$

for each compact K in D; see Lemma 2 of [3], the crucial Lemma there. It follows from (26) that $Q_t^D g$ converges uniformly to g in each K, and so g is continuous in D.

Since D is regular, $g = 1$ on D^c. It remains to prove that as $x \in D$, $x \to z \in \partial D$, $g(x)$ converges to 1. Indeed

$$g(x) - 1 = E^x\{e(\tau_D) - 1; \tau_D \le t\} + E^x\{e(\tau_D) - 1; t < \tau_D\};$$
$$|g(x) - 1| \le E^x\{e_{A^*}(t) - 1\} + E^x\{|e(t)g(X_t) - 1|; t < \tau_D\}.$$

The first term tends to zero as $t \to 0$ by Lemma 3 applied to A^*. For the second term we have

$$E^x\{|e(t)g(X_t) - 1|; t < \tau_D\} \le CE^x\{e^2(t)\}^{1/2}(P^x(t < \tau_D))^{1/2} + P^x\{t < \tau_D\}$$

since $g \le C$, $\sup_x \sup_{0\le t\le 1} E^x\{e(t)^2\} < \infty$, and $P^x(t < \tau_D) \to 0$ as $x \to \partial D$, the second term also tends to zero.

The following complement to the gauge theorem is easy but important. It does not depend on the latter and holds for any open set D.

PROPOSITION 8. $\inf_{x\in E} g(x) > 0$.

PROOF. By Jensen's inequality

$$g(x) \ge E^x\{e^{-A^{(2)}(\tau_D)}\} \ge e^{-E^x\{A^{(2)}(\tau_D)\}}.$$

Since

$$\sup_{x\in E} E^x\{A^{(2)}(\tau_D)\} \le \sup_{x\in E} V^{(2)}(x) < \infty,$$

the proposition follows.

4. Consequences of bounded gauge. In this section we assume that the gauge for (D, A) is bounded, when D is a nonempty open set, not necessarily relatively compact. We shall deduce several consequences from this, some of which are strictly stronger than the initial assumptions. Note that the boundedness of the gauge for (D, A) does not imply that for (D, A^*). Nevertheless we have the following useful result.

THEOREM 4. *We have*

$$\sup_{x\in E} E^x\left\{\int_0^{\tau_D} e_A(t)\,dA_t^*\right\} < \infty. \tag{27}$$

PROOF. We begin with the calculus formula

$$\int_0^s e^{A^*(t)}\,dA^*(t) = e^{A^*(s)} - 1.$$

Since $A(t) \le A^*(t)$ it follows that

$$E^x\left\{\int_0^t e_A(t)dA^*(t)\right\} \le E^x\{e_{A^*}(t)\} - 1 \le C(t) \tag{28}$$

where $C(t)$ is a constant depending on t but not x, by Lemma 3 applied to A^*. Now we have, writing τ for τ_D and $e(t)$ for $e_A(t)$ below,

$$\begin{aligned} E^x\left\{\int_0^\tau e(t)\,dA_t^*\right\} &= E^x\left\{\sum_{n=0}^\infty \int_{\tau\wedge n}^{\tau\wedge(n+1)} e(t)\,dA^*(t)\right\} \\ &= \sum_{n=0}^\infty E^x\left\{n<\tau; \int_n^{\tau\wedge(n+1)} e(t)\,dA_t^*\right\} \\ &= \sum_{n=0}^\infty E^x\left\{n<\tau; e(n)E^{X_n}\left\{\int_0^{\tau\wedge 1} e(t)\,dA_t^*\right\}\right\} \\ &\le C(1)\sum_{n=0}^\infty E^x\{n<\tau; e(n)\} \end{aligned} \tag{29}$$

by (28). The convergence of the series in (29) is proved in the following lemma.

LEMMA 9. *For every $\delta > 0$, there exists a constant $C(\delta)$ such that*

$$\|g\|_\infty^{-1} g(x) \le \sum_{n=0}^\infty E^x\{n\delta < \tau; e(n\delta)\} \le C(\delta).$$

PROOF. Since the proof is the same for any $\delta > 0$ we take $\delta = 1$. For each $x \in D$ the quantity $E^x\{n < \tau; e(\tau)\}$ decreases to zero. Hence for any $\varepsilon > 0$, we can choose N and a set F such that $\sup_{x\in F} E^x\{N < \tau; e(\tau)\} < \varepsilon$ and $m(D \setminus F) < \varepsilon$. By choosing a compact subset of F if necessary we may assume that F itself is compact. Here ε is such that $\sup_x E^x\{(e(\tau_1))^2\} \le 2$ where $\tau_1 = \tau_{D\setminus F}$. This is possible by the remark immediately after the proof of Lemma 4. We have

$$\begin{aligned} E^x\{2N<\tau; e(\tau)\} &\le E^x\{N<\tau_1; e(\tau)\} + E^x\{\tau_1 \le N, 2N < \tau; e(\tau)\} \\ &\le \|g\|_\infty E^x\{N<\tau_1; e(\tau_1)\} + E^x\{\tau_1<\tau; e(\tau_1)E^{X(\tau_1)}[N<\tau; e(\tau)]\}. \end{aligned}$$

The first term above can be estimated by $E^x\{e^2(\tau_1)\}^{1/2}P^x(\tau_1 > N)^{1/2}$ which is uniformly small if N is large, by Corollary to Lemma 2. On the set $\{\tau_1 < \tau\}$, $X_{\tau_1} \in F$ so that the second is bounded by 2ε. Thus for large N,

$$\sup_x E^x\{N<\tau; e(\tau)\}$$

is small. Now from Proposition 8, g is bounded from below, say $g \ge m$. Then

$$\begin{aligned} E^x\{N<\tau; e(N)\} &\le \frac{1}{m}E^x\{N<\tau; e(N)g(X_N)\} \\ &\le \frac{1}{m}E^x\{N<\tau; e(\tau)\}. \end{aligned}$$

Therefore we see that for N large enough

$$\sup_x E^x\{N<\tau; e(N)\} \le \alpha < 1.$$

By the Markov property for any k:

$$\sup_x E^x\{kN < \tau; e(kN)\} \le \alpha^k.$$

Further if $j < N$:

$$\begin{aligned} E^x\{kN + j < \tau; e(kN + j)\} &= E^x\{j < \tau; e(j) E^{X_j}\{kN < \tau; e(kN)\}\} \\ &\le \alpha^k \sup_{x, j \le N} E^x\{e(j)\} = \alpha^k M. \end{aligned}$$

All these inequalities give

$$\begin{aligned} \sum_n E^x\{n < \tau; e(n)\} &= \sum_{k=0}^{\infty} \sum_{j=1}^{N} E^x\{kN + j < \tau; e(kN + j)\} \\ &\le NM \sum_0^{\infty} \alpha^k = \frac{NM}{1 - \alpha} = C(1). \end{aligned}$$

Finally we have

$$\begin{aligned} g(x) &= \sum E^x[n\delta < \tau < (n+1)\delta; e(\tau)] \\ &\le \sum E^x[n\delta < \tau; e(n\delta) g(X_{n\delta})] \le \|g\|_\infty \sum E^x[n\delta < \tau; e(n\delta)] \end{aligned}$$

which proves the result.

The following consequence is stronger than the gauge theorem.

THEOREM 5.

$$\sup_{x \in D} E^x \left\{ \sup_{0 \le t \le \tau_D} e_A(t) \right\} < \infty. \tag{31}$$

PROOF. We have

$$e^{A(t)} = 1 + \int_0^t e_A(s)\, dA(s) \le 1 + \int_0^t e_A(s)\, dA^*(s),$$

hence

$$\sup_{0 \le t \le \tau_D} e_A(t) \le 1 + \int_0^{\tau_D} e_A(s)\, dA^*(s)$$

and (31) follows from (27).

COROLLARY.

$$g(x) = E^x \left[\int_0^{\tau_D} e(t)\, dA_t \right] + 1.$$

PROOF. We can use Fubini by (27).

Before we proceed let us state several properties that are equivalent to the boundedness of the gauge. These are similar to Theorem 3.2 in [8] but somewhat improved. Recall the notation Q_t^D from §3.

THEOREM 6. *Any one of the following propositions is equivalent to the boundedness of the gauge.*

(i) *For some $\delta > 0$ and some x,*

$$\sum_n Q_{n\delta}^D 1(x) < \infty. \tag{32}$$

(ii) *For all* $\delta > 0$, *and all* x, (32) *is true.*
(iii) *For some* x,

$$\int_0^\infty Q_t^D 1(x)\,dt < \infty.$$

(iv) *There exists* t *such that*

$$\sup_{x\in D} Q_t^D 1(x) < 1.$$

(v) *There exists* $C > 0$ *and* $b > 0$ *such that*

$$\sup_{x\in D} Q_t^D 1(x) \le Ce^{-bt}.$$

PROOF. That (iv) and (v) are equivalent is by the usual use of the Markov property and (v) clearly implies (i). Of course (ii) implies (i). In short, the following equivalences are clear or easy:

$$\text{(iv)} \Leftrightarrow \text{(v)};\ \text{(v)} \Rightarrow \text{(i)};\ \text{(ii)} \Rightarrow \text{(i)}.$$

If (i) is valid then $g(x) < \infty$ and hence bounded by the Gauge Theorem. By Lemma 9 then (ii) is valid and in the course of the proof of Lemma 9 we also saw that (iv) is valid. We now show that (i) is equivalent to (iii).

Fix x and suppose (32) is true. We have

$$\begin{aligned}\int_0^\infty Q_t 1(x)\,dt &= \sum_0^\infty \int_{n\delta}^{n\delta+\delta} Q_t 1(x)\,dt = \sum_0^\infty \int_{n\delta}^{n\delta+\delta} E^x[t<\tau; e(t)]\,dt \\ &= \sum_0^\infty \int_{n\delta}^{n\delta+\delta} E^x[n\delta<\tau; e(n\delta)E^{X(n\delta)}[t-n\delta<\tau; e(t-n\delta)]]\,dt \\ &\le C\sum_0^\infty E^x[n\delta<\tau; e(n\delta)] < \infty\end{aligned}$$

where $C = \sup_x E^x\{e_{A^*}(1)\}$.

Suppose now that (iii) holds. For $n\delta < t < (n+1)\delta$ we have

$$\begin{aligned}Q_{(n+1)\delta}1(x) &= E^x[(n+1)\delta<\tau; e(n\delta+\delta)] \le E^x[t<\tau; e(n\delta+\delta)] \\ &= E^x[t<\tau; e(t)E^{X_t}[e((n+1)\delta-t)]] \le CE^x[t<\tau; e(t)]\end{aligned}$$

where C is as before. Hence

$$Q_{n\delta+\delta}1(x) \le C\delta^{-1}\int_{n\delta}^{n\delta+\delta} Q_t 1(x)\,dt$$

and the quantity in (i) is bounded by $1 + C\delta^{-1}\int_0^\infty Q_t 1(x)\,dt$. Thus all five are equivalent to the boundedness of the gauge.

5. The super gauge theorem. This is another consequence of the boundedness of the gauge, which asserts that if the gauge for (D, A) is bounded then so is it for $(D, (1+\varepsilon)A)$ for sufficiently small $\varepsilon > 0$. This phenomenon has an intuitive base in the old case, having to do with eigenvalues (spectrum) of an associated Schrödinger operator. See §6. We shall give two distinct proofs of this interesting result. The first is more direct and similar to the old treatment. The second relates it to the recent development in "bounded mean oscillations (BMO)" and is included here on account of this amusing connection.

THEOREM 7. *If the gauge is bounded (in E) then there exists $\varepsilon > 0$ such that*

$$\sup_{x\in E} E^x\{e(\tau_D)^{1+\varepsilon}\} < \infty. \tag{33}$$

PROOF. By Theorem 6, there exists t such that for all $x \in D$,

$$Q_t^D 1(x) = E^x\{t < \tau_D; e(t)\} < 1. \tag{34}$$

Consider

$$\begin{aligned} E^x\{|e(t)^{1+\varepsilon} - e(t)|\} &= E^x\{e(t)|e(t)^\varepsilon - 1|\} \\ &\le E^x\{e(t)^2\}^{1/2} E^x\{|e^{\varepsilon A(t)} - 1|^2\}^{1/2} \\ &\le C(t) E^x\{|e^{\varepsilon A(t)} - 1|^2\}^{1/2}, \end{aligned} \tag{35}$$

where $C(t)$ is a constant independent of x, by Lemma 3 applied to $2A$. We have

$$|e^{\varepsilon A(t)} - 1| \le e^{\varepsilon A^*(t)} - 1.$$

Now since A^* is increasing

$$e^{\varepsilon A^*(t)} - 1 = \int_0^t e^{\varepsilon A^*(s)} \varepsilon\, dA^*(s) \le \varepsilon e^{\varepsilon A^*(t)} A^*(t).$$

Hence

$$E^x\{(e^{\varepsilon A^*(t)} - 1)^2\} \le \varepsilon^2 E^x\{e^{2\varepsilon A^*(t)} A^*(t)^2\} \le \varepsilon^2 C(t)$$

by Lemma 3. It follows from this and (35) that as $\varepsilon \downarrow 0$, $E^x\{|e(t)^{1+\varepsilon} - e(t)|\}$ converges to zero uniformly in $x \in E$. Therefore by (34), if ε is sufficiently small

$$\sup_{x\in E} E^x\{t < \tau_D; e(t)^{1+\varepsilon}\} < 1.$$

But $e_A(t)^{1+\varepsilon} = e_{(1+\varepsilon)A}(t)$. Hence by Theorem 6(iv) applied to $(1+\varepsilon)A$ we conclude that the gauge for $(D, (1+\varepsilon)A)$ is bounded.

For the second proof of Theorem 7 we introduce the probability measure, for $\Lambda \in \mathcal{F}_\infty$,

$$M^x(\Lambda) = \frac{1}{g(x)} E^x\{e_A(\tau_D) 1_\Lambda\}, \qquad x \in E, \tag{36}$$

where g is the gauge for (D, A). This is well defined since $g > 0$ by Proposition 8, and $M^x(Q) = 1$ since $g(x) < \infty$. Now we compute the conditional expectation. Put $\mathcal{F}_t^\tau = \mathcal{F}_t 1_{(t<\tau)}$ where $\tau = \tau_D$,

$$\begin{aligned} M^x\{A_\tau^* - A_t^* | \mathcal{F}_t\} &= M^x\{A_\tau^*(\theta_t) | \mathcal{F}_t^\tau\} \\ &= \frac{1}{g(x)} E^x\{e(t)(e(\tau)A_\tau^*)(\theta_t) | \mathcal{F}_t^\tau\} \\ &= \frac{1}{g(x)} E^x\{t < \tau; e(t) E^{X_t}\{e(\tau)A_\tau^*\} | \mathcal{F}_t^\tau\}. \end{aligned} \tag{37}$$

But for all $x \in D$,

$$\begin{aligned} E^x\{e(\tau)A_\tau^*\} &= E^x\left\{\int_0^\tau e(\tau)\, dA_t^*\right\} = E^x\left\{\int_0^\tau e(t) E^{X_t}\{e(\tau)\}\, dA_t^*\right\} \\ &\le \|g\|_\infty E^x\left\{\int_0^\tau e(t)\, dA_t^*\right\} \le C_1 \end{aligned}$$

by Theorem 4. Hence the quantity in (37) is bounded by

$$\frac{1}{g(x)}E^x\{t<\tau;e(t)\}C_1 \le \frac{1}{C_2}E^x\left\{\sup_{0\le t\le\tau} e(t)\right\}C_1$$

where C_2 is a lower bound for g by Proposition 8. By Theorem 5, the last quantity is bounded in D. Thus there exists a constant C such that for all $x \in D$,

$$M^x\{A^*_\tau - A^*_t|\mathcal{F}^\tau_t\} < C.$$

Now replace A^* by εA^* above, then for $\varepsilon < C^{-1}$, we obtain

$$M^x\{\varepsilon(A^*_\tau - A^*_t)|\mathcal{F}^\tau_t\} < \theta < 1.$$

It follows by Theorem 109 (John-Nirenberg) in [**9**, p. 193] that

$$E^x\{e_{(1+\varepsilon)A}(\tau)\} = E^x\{e(\tau)e_{\varepsilon A^*}(\tau)\} = g(x)M^x\{e^{\varepsilon A^*_\tau}\} < \infty.$$

A curious consequence of the super gauge theorem is that either there exists $\varepsilon_0 > 0$ such that the gauge for $(D,(1+\varepsilon)A)$ is bounded for all $\varepsilon < \varepsilon_0$, but infinite for $\varepsilon = \varepsilon_0$, or it is bounded for (D, pA) for all $p > 0$. The latter holds e.g. if $A = -A^{(2)}$. Let us add that by the super gauge theorem, all previous results such as Theorems 4 and 5 are valid when the A there is replaced by $(1+\varepsilon)A$ for sufficiently small $\varepsilon > 0$. For instance,

$$\sup_{x\in D} E^x\left\{\sup_{0\le t\le \tau_0} e_A(t)^{1+\varepsilon}\right\} < \infty.$$

An interesting application of the super gauge theorem will be given in §6. For another application in the old case see Chung [**4**].

6. Green potentials. We call the function

$$U^D f(x) = E^x\left\{\int_0^{\tau_D} f(X_t)\,dt\right\} = \int_0^\infty P^D_t f(x)\,dt, \qquad f \in L^\infty(D), \tag{38}$$

the Green potential of f. Indeed if X is the Brownian motion in R^d, $d \ge 1$, this is the classical Green potential. Here we need D to be relatively compact to ensure that $\tau_D < \infty$ a.s.; otherwise we have to consider $U^{D,\lambda}$ for $\lambda > 0$ as in §1. For the sake of simplicity we shall assume (for the first time in this paper) that X is transient in the sense that for each compact subset K of E,

$$x \to U(x,K) = E^x\left\{\int_0^\infty 1_K(X_t)\,dt\right\} \tag{39}$$

is a bounded function in E. Then it follows that for all relatively compact open sets D $\sup_{x\in E} E^x\{\tau_D\} < \infty$, so that the Green potential in (38) is bounded in E. Now, with the additive functional A specified above, we generalize the notion as follows:

$$V^D f(x) = E^x\left\{\int_0^{\tau_D} e_A(t)f(X_t)\,dt\right\} = \int_0^\infty Q^D_t f(x)\,dt. \tag{40}$$

[Note that this expression is given in Theorem 6, (iii).] This will be called the A-Green function, so that U^D becomes the O-Green potential. There is a simple relation between these, given in the old case in Chung [**5**]. It is easily extended to the present more general case; but we shall spell out half of it below.

THEOREM 8. *Assume the gauge is bounded. We have*

$$\tag{41} \begin{aligned} V^D f &= U^D f + E^{\cdot}\left\{\int_0^{\tau_D} e_A(s) U^D f(X_s)\,dA_s\right\} \\ &= U^D f + E^{\cdot}\left\{\int_0^{\tau_D} V^D f(X_s)\,dA_s\right\}. \end{aligned}$$

PROOF. We may assume $f \in L_+^\infty(D)$. As just remarked, $U^D f$ is bounded. We have, writing τ for τ_D, $e(t)$ for $e_A(t)$:

$$\tag{42} \begin{aligned} E^x\left\{\int_0^\tau e(s) U^D f(X_s)\,dA_s\right\} &= E^x\left\{\int_0^\tau e(s) E^{X_s}\left\{\int_0^\tau f(X_r)\,dr\right\} dA_s\right\} \\ &= E^x\left\{\int_0^\tau e(s)\int_s^\tau f(X_r)\,dr\,dA_s\right\}. \end{aligned}$$

In order to reverse the order of integration above, we note that $|dA_s| \le dA_s^*$, and

$$E^x\left\{\int_0^\tau e(s)\int_s^\tau f(X_r)\,dr\,dA_s^*\right\} = E^x\left\{\int_0^\tau e(s) U^D f(X_s)\,dA_s^*\right\} < \infty$$

by Theorem 4. Therefore, we can reverse the order of integration in (42) and obtain

$$\begin{aligned} E^x\left\{\int_0^\tau f(X_r)\int_0^\tau e(s)\,dA_s\,dr\right\} &= E^x\left\{\int_0^\tau f(X_r)[e(r)-1]\,dr\right\} \\ &= V^D f(x) - U^D f(x). \end{aligned}$$

This establishes the first relation in (41); the second is similar.

Theorem 8 is the integrated form of relations between the infinitesimal generators of the semigroups (P_t^D) and (Q_t^D). In the old case, when X is the Brownian motion, and

$$A(t) = \int_0^t q(X_s)\,ds$$

for an appropriate class of functions q, these are the Laplacian $\Delta/2$ and Schrödinger operators $\Delta/2 + q$, respectively, and $(\Delta/2 + q)V^D f = -f$.

In the general case recalling (6) and (7) of §1, we define the Revuz measure (see [**10**]) associated with $U_A 1$ as follows: for $\phi \in L^\infty$

$$\mu_A(\phi) = \lim_{t\downarrow 0} t^{-1}\int_{\mathcal{E}} E^x\left\{\int_0^t \phi(X_s)\,dA_s\right\} m(dx)$$

where m is a reference measure. Then the Revuz measure associated with the $U_A f$ in (7) is given by $f \cdot \mu_A$. In the old case where $A(t) = \int_0^t q(X_s)\,ds$ we have $\mu_A = q \cdot m$. Now put $u = V^D f$, and

$$B_t = \int_0^t f(X_s^D)\,ds + \int_0^t u(X_s^D)\,dA_s.$$

Then by the second relation in (41), we have

$$u(x) = E^x\{B_\infty\}, \qquad x \in D.$$

Hence the Revuz measure associated with u is given by

$$\tag{43} \mu_B = f \cdot m + u \cdot \mu_A.$$

In the old case this reduces to

$$\mu_B = (f + uq) \cdot m.$$

Since in that case q is also given by

$$q = (-2^{-1}\Delta)U_A 1 \tag{44}$$

this shows

$$-2^{-1}\Delta u = f + uq$$

or

$$(2^{-1}\Delta + q)u = -f; \tag{45}$$

namely that u is a solution of the nonhomogeneous Schrödinger equation above. In the cases where the Laplacian operator in (44) is replaced by a more general differential operator, an equation similar to (45) results. Moreover, the old case may be slightly generalized so that the measure $q \cdot m$ becomes a Radon measure ν. That is an easy consequence of our general theorem applied to the additive functional generated by $U\nu$.

A remarkable connection between the gauge and A-Green potential will now be discussed. Consider the functions for $f \in L^\infty_+(D)$ and $x \in D$,

$$\begin{aligned}\phi(x) &= E^x\left\{e(\tau_D)\int_0^{\tau_D} f(X_t)\,dt\right\}\\ &= E^x\left\{\int_0^{\tau_D} f(X_t)E^{X_t}[e(\tau_D)]\,dt\right\} = E^x\left\{\int_0^{\tau_D} e(t)g(X_t)f(X_t)\,dt\right\}.\end{aligned}$$

It follows at once by Proposition 8 that when the gauge is bounded $\phi(x) \sim V^D f(x)$ where for two positive functions ϕ_1 and ϕ_2 we use the notation $\phi_1 \sim \phi_2$ to denote that ϕ_1/ϕ_2 is both bounded above and away from zero.

Next by the super gauge theorem, for sufficiently small ε we have

$$\sup_x E^x\{e(\tau_D)^{1+\varepsilon}\} \le M_1 < \infty.$$

Now by Hölder's inequality,

$$\phi(x) \le E^x\{e(\tau_D)^{1+\varepsilon}\}^{1/(1+\varepsilon)} E^x\left\{\left(\int_0^{\tau_D} f(X_t)\,dt\right)^{(1+\varepsilon)/\varepsilon}\right\}^{\varepsilon/(1+\varepsilon)}.$$

We may choose ε so that $(1+\varepsilon)/\varepsilon = p$ is an even integer. Then by an argument we have used before in the proof of Lemma 3, if

$$\sup_{x\in E} U^D f(x) = M_2 < \infty,$$

we have

$$E^x\left\{\left(\int_0^{\tau_D} f(X_t)\,dt\right)^p\right\} \le p!M_2^p.$$

Consequently there exists M:

$$\sup_{x\in E}\phi(x) \le M \sup_{x\in E} U^D f(x), \tag{46}$$

where $M = M_1(p!)^{1/p}$. (M_1 depends on p.)

In the old case when X is the Brownian motion and

$$A(t) = \int_0^t q(X_s)\, ds$$

and when ∂D is smooth, say of the class $C^{1,1}$, Zhao [6] proved that there exists M such that for all x, $\phi(x) \le MU^D f(x)$. The result above is a weaker version of this but valid in our general case without any assumption on ∂D.

REFERENCES

1. R. M. Blumenthal and R. K. Getoor, *Markov processes and potential theory*, Academic Press, New York, 1968.
2. K. L. Chung, *Lectures from Markov processes to Brownian motion*, Springer-Verlag, Berlin and New York, 1985.
3. ——, *Doubly Feller process with multiplicative functional*, Seminar on Stochastic Processes, Birkhauser, 1985, pp. 63–78.
4. ——, *Properties of finite gauge with an application to local time*, Essays in honor of Carl-Gustav Esseen, Uppsala, Sweden, 1983, pp. 16–24.
5. ——, *Notes on the inhomogeneous Schrödinger equation*, Seminar on Stochastic Processes, Birkhäuser, 1984, pp. 55–62.
6. K. L. Chung and Z. Zhao, forthcoming monograph.
7. K. L. Chung and M. Rao, *A new setting for potential theory*, Ann. Inst. Fourier (Grenoble) **30** (1980), 167–198.
8. ——, *Feynman-Kac functional and the Schrödinger equation*, Seminar on Stochastic Processes, Birkhäuser, 1981, pp. 1–29.
9. C. Dellacheri and P. A. Meyer, *Probabilités et potentiel*, Hermann, Paris, 1980.
10. D. Revuz, *Mesures associees aux fonctionelles additives de Markov.* I, Trans. Amer. Math. Soc. **148** (1970), 501–531.

DEPARTMENT OF MATHEMATICS, STANFORD UNIVERSITY, STANFORD, CALIFORNIA 94305

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF FLORIDA, GAINESVILLE, FLORIDA 32611

Reprinted from *Seminar on Stochastic Processes, 1990*, Birkhäuser, 1991, pp. 5–14.
With kind permission of Springer Science and Business Media.

Paul Lévy's Way to His Local Time

A.A. BALKEMA and K.L. CHUNG

0. Foreword by Chung

In his 1939 paper [1] Lévy introduced the notion of local time for Brownian motion. He gave several equivalent definitions, and towards the end of that long paper he proved the following result. Let $\epsilon > 0$, $t > 0$, $B(0) = 0$,

$$L_\epsilon(t) = m\{s \in [0,t] \mid 0 < B(s) < \epsilon\}/\epsilon \tag{0.1}$$

where $B(t)$ is the Brownian motion in R and m is the Lebesgue measure. Then almost surely the limit below exists for all $t > 0$:

$$\lim_{\epsilon \to 0} L_\epsilon(t) = L(t). \tag{0.2}$$

This process $L(.)$ is Lévy's local time.

As I pointed out in my paper which was dedicated to the memory of Lévy, [2; p.174], there is a mistake in the proof given in [1], in that the moments of occupation time for an excursion were confounded with something else, not specified. Apart from this mistake which I was able to rectify in Theorem 9 of [2], Lévy's arguments can (easily) be made rigorous by standard "bookkeeping". As any serious reader of Lévy's work should know, this is quite usual with his intensely intuitive style of writing. Hence at the time when I wrote [2], I did not deem it necessary to reproduce the details. Nevertheless I scribbled a memorandum for my own file. Later, after I lectured on the subject in Amsterdam in 1975, I sent that memo to Balkema in the expectation that he would render it legible. This

valuable sheet of paper has apparently been lost. In my reminiscences of Lévy [3], spoken at the Ecole Polytechnique in June, 1987, I recounted his invention of local time and the original proof of the theorem cited above. It struck me as rather odd that although a supposedly historical account of this topic was given in Volume 4 of Dellacherie-Meyer's encyclopaedic work [4], Lévy's 1939 paper was not even listed in the bibliography. This must be due to the failure of the authors to realize that the contents of that paper were not entirely reproduced in Lévy's 1948 book [5]. Be that as it may, incredible events posterior to the Lévy conference in 1987 (see the Postscript in [3]) have convinced me that very few people have read, much less understood, Lévy's own way to his invention. I have therefore asked Balkema to write a belated exposition based on my 1975 lectures on Brownian motion. Together with the results in my paper [2] on Brownian excursions this forms the basis of the present exposition of Lévy's ideas about local time. Now I wonder who among the latter-day experts on local time will have the curiosity (and humility) to read it?

1. Local time of the zero set of Brownian motion

One of the most striking results on Brownian motion is Lévy's formula:

$$B \stackrel{\mathrm{d}}{=} |B| - L^*$$

where B is Brownian motion and L^* is the local time of $|B|$ in zero defined in terms of the zero set of B. Lévy considered the pair $(M - B, M)$ where M is the max process for Brownian motion:

$$M_t = \max\{B(s) \mid s \leq t\},$$

and proved that the process $Y = M - B$ is distributed like the process $|B|$, using the at that time not yet rigorously established strong Markov property for Brownian motion. In one picture we have the continuous increasing process M and dangling down from it the process Y (distributed like $|B|$). Note that

M increases only on the zero set of Y. Problem: Can one express the sample functions of the increasing process M in terms of the sample functions of the process Y?

Let us define

$$T_u = \inf\{t > 0 \mid M(t) > u\} \qquad u \geq 0.$$

This is the right-continuous inverse process to M. Lévy observed that it is a pure jump process with stationary independent increments. It has Lévy measure $\rho(y, \infty) = \sqrt{(2/\pi y)}$ on $(0, \infty)$. There is a 1-1 correspondence between excursion intervals of Y and jumps of the Lévy process T. Hence the number of excursions of Y in $[0, T_u]$ of duration $> c$ is equal to the number $N = N_c(u)$ of jumps of T of height $> c$ during the interval $[0, u]$. For a Lévy process this number is Poisson distributed with parameter $u\rho(c, \infty) = u\sqrt{(2/\pi c)}$ in our case. In fact if we keep u fixed then $t \to N_{c(t)}$, with $c(t) = 2/\pi t^2$, is the standard cumulative Poisson process on $[0, \infty)$ with intensity u. The strong law of large numbers (for exponential variables) implies

$$N_c(u)/\sqrt{(2/\pi c)} \to u \text{ a.s.} \qquad \text{as } c = c(t) \to 0. \tag{1.1}$$

Now vary u. The counting process $N_c : [0, \infty) \to 0, 1, \ldots$ will satisfy (1.1) for all rational $u \geq 0$ for all ω outside some null set Ω_0 in the underlying probability space. For these realizations we have weak convergence of monotone functions and hence uniform convergence on bounded subsets (since the limit function is continuous). In particular we have convergence for each $u \geq 0$, also if $u = M_t(\omega)$ depends on ω. This proves:

Theorem 1.1 (Lévy). Let B be a Brownian motion and let $N_c^*(t)$ denote the number of excursion intervals of length $> c$ contained in $[0, t]$. Then

$$N_c^*(t)/\sqrt{(2/\pi c)} \to L^*(t) \text{ a.s.} \qquad \text{as } c \to 0$$

for some process L^* with continuous increasing sample paths in the sense of weak convergence. Moreover $(|B|, L^*) \stackrel{\mathrm{d}}{=} (M - B, M)$.

Corollary. L^* is unbounded a.s. and $L^*(0) = 0$.

Note that local time L^* has been defined in terms of the zero set $Z = \{t \geq 0 \mid B(t) = 0\}$. We call this process L^* the local time of the zero set of Brownian motion in order to distinguish it from the process L introduced in (0.2). The process L_ϵ in (0.1) depends on the behaviour of Brownian motion in the ϵ-interval $(0, \epsilon)$. For a discussion of local times for random sets see Kingman [6]. Here we only observe that one can construct another variant of local time in 0 by counting excursions of sup norm $> c$ rather than excursions of duration $> c$. The Lévy measure then is dy/y^2 rather than $dy/\sqrt{(2\pi y^3)}$. This latter procedure has the nice property that it is invariant for time change and hence works for any continuous local martingale.

The next result essentially is an alternative formulation of Theorem 1.1.

Lemma 1.2. Let $u > 0$, and let U_c be the upper endpoint of the $K(c)$th excursion of the Brownian motion B of duration $> c$. Assume that a.s. $K(c) \sim u\sqrt{(2/\pi c)}$ as $c \to 0$. Then $U_c \to T^*_u$ a.s. as $c \to 0$, where

$$T^*_u(\omega) = \inf\{t > 0 \mid L^*_t(\omega) > u\}. \tag{1.2}$$

Proof. The process $u \mapsto T^*_u$ is a Lévy process since $L^* \stackrel{d}{=} M$ by Theorem 1.1. Hence it has no fixed discontinuities. Choose a sample point ω in the underlying probability space such that

1) the function $L^*(\omega)$ in Theorem 1.1 is continuous, increasing, unbounded, and vanishes in $t = 0$,

2) the limit relation of Theorem 1.1 holds,

3) $K(c)(\omega) \sim u\sqrt{(2/\pi c)}, \quad c \to 0,$

4) the function $T^*(\omega)$ is continuous at the point u.

We omit the symbol ω in the expressions below. Let $0 < u_1 < u < u_2$ and let $N_i(c)$ denote the number of excursions of length $> c$ in the interval $[0, T^*_{u_i}]$ for $i = 1, 2$. Theorem 1.1 gives the asymptotic relation $N_i(c) \sim u_i\sqrt{(2/\pi c)}$ as $c \to 0$.

Hence for all sufficiently small c we have the inequality $N_1(c) < K(c) < N_2(c)$, and therefore $T^*_{u_1} < U_c < T^*_{u_2}$. The continuity of the sample function T^* at u then implies that $U_c \to T^*_u$.

This innocuous-looking lemma enables us to consider the $S(c)$ in Section 2 with a constant $n(c)$, rather than a random number, which would entail subtle considerations of the dependence between the sequence $\{\psi_n\}$ and the process L^*.

2. Local time as a limit of occupation time

In order to prove Theorem 1.1 using the occupation time of the interval $(0, \epsilon)$, $\epsilon \to 0$, rather than the number of excursions, one needs a bound on the second moment of the occupation time of the interval $(0, \epsilon)$ for the excursions. We begin with a simple but fundamental result.

Theorem 2.1. For fixed $c > 0$ the sequence of excursions of Brownian motion of duration exceeding c is i.i.d. provided the excursions are shifted so as to start in $t = 0$.

Proof. The upper endpoint τ_1 of the first excursion φ_1 of duration $> c$ is optional. By the strong Markov property the process $B_1(t) = B(\tau_1 + t)$, $t \geq 0$, is a Brownian motion and is independent of φ_1. Hence φ_1 is independent of the sequence $(\varphi_2, \varphi_3, \ldots)$ and $\varphi_1 \stackrel{d}{=} \varphi_2$ since φ_2 is the first excursion of B_1 of duration $> c$. Now proceed by induction.

As an aside let us show, as Lévy did, that this theorem by itself gives local time up to a multiplicative constant: Choose a sequence c_n decreasing to zero. We obtain an increasing family of i.i.d. sequences of excursions which contains all the excursions of Brownian motion. Each of these i.i.d. sequences acts as a clock. The large excursions of duration $> c_0$ ring the hours. The next sequence contains all excursions of duration $> c_1$ and ticks off the minutes. The next one the seconds, etc. Note that the number of minutes per hour is random: The sequence of excursions of duration $> c_1$ is i.i.d. and hence the subsequence of excursions of

duration $> c_0$ is generated by a selection procedure which gives negative binomial waiting times with expectation $\sqrt{(c_0/c_1)}$. Similarly the number of seconds per hour is negative binomial with expectation $\sqrt{(c_0/c_2)}$. If we standardize the clocks so that the intertick times of the nth clock are $\sqrt{(c_n/c_0)}$ then the clocks become ever more accurate. The limit is local time for Brownian motion. Pursuing this line of thought one can show that the excursions of Brownian motion form a time homogeneous Poisson point process on a product space $[0,\infty)\times E$ where E is the space of continuous excursions and the horizontal axis is parametrized by local time. See Greenwood and Pitman [7] for details.

We now return to our main theme. Let $\psi_1, \psi_2, \ldots$ be the i.i.d. sequence of *positive* excursions of duration $> c$. This is a subsequence of the sequence (φ_n) of theorem 2.1. Given $\epsilon > 0$ let $f_\epsilon(\psi_n)$ denote the occupation time of the space interval $(0, \epsilon)$ for the nth excursion ψ_n:

$$f_\epsilon(\psi_n) = m\{t > 0 \mid 0 < \psi_n(t) < \epsilon\}$$

and set

$$X_n = f_\epsilon(\psi_n)/\epsilon.$$

Section 3 contains the proofs of the following key estimates:

$$\mathcal{E}(X_n) \sim \sqrt{2\pi c} \qquad c \to 0 \tag{2.1}$$

$$\mathcal{E}(X_n^2) \le 6\epsilon\sqrt{c} \qquad 0 < \epsilon, 0 < c. \tag{2.2}$$

Now define

$$Y_n = X_n - \mathcal{E}X_n$$

$$S(c) = Y_1 + \cdots + Y_{n(c)}$$

where $n(c) = [u/\sqrt{2\pi c}]$ for some fixed $u > 0$. We are interested in the case $c \to 0$. We have by (2.2)

$$\mathcal{E}(Y_n^2) = \sigma^2(X_n) \le 6\epsilon\sqrt{c}$$

which gives

$$\mathcal{E}(S(c)^2) \leq 6\epsilon u.$$

By (2.1) we have

$$\mathcal{E}(X_1 + \cdots + X_{n(c)}) = n(c)\mathcal{E}X_1 \to u \quad \text{as } c \to 0.$$

Let U_c denote the upper endpoint of the $n(c)$th positive excursion $\psi_{n(c)}$. Note that $\psi_{n(c)} = \varphi_{K(c)}$ is the $K(c)$th excursion of duration exceeding c and that $K(c) \sim 2n(c)$ a.s. by the strong law of large numbers for a fair coin. Lemma 1.2 shows that $U_c \to T_u^*$ a.s. as $c \to 0$ where T_u^* is defined in (1.2). Hence

$$X_1 + \cdots + X_{n(c)} \to L_\epsilon(T_u^*) \text{ a.s.} \quad \text{as } c \to 0. \tag{2.3}$$

Fatou' s lemma then yields

Lemma 2.2. $\mathcal{E}(L_\epsilon(T_u^*) - u)^2 \leq \liminf_{c\to 0} \mathcal{E}(S(c)^2) \leq 6\epsilon u.$

This inequality will enable us to prove (0.2).

Theorem 2.3. Define L_ϵ by (0.1). Then

$$L_\epsilon(t) \to L^*(t) \text{ a.s.} \quad \text{as } \epsilon \to 0 \tag{2.4}$$

in the sense of weak convergence of monotone functions.

Proof. It suffices to show that for each rational $u > 0$ the scaled occupation time

$$L_\epsilon(T_u^*) = m\{t \in [0, T_u^*] \mid 0 < B(t) < \epsilon\}/\epsilon \to u \text{ a.s.} \quad \text{as } \epsilon \to 0.$$

Since occupation time is increasing for fixed $\epsilon > 0$ and local time is continuous this will imply weak convergence. In the definition of $L_\epsilon(t)$ as a ratio both numerator and denominator are increasing in ϵ. Hence it suffices to prove the convergence for $\epsilon_n = n^{-4}$, as $n \to \infty$. We have by Lemma 2.2

$$p_n = P\{|L_{\epsilon_n}(T_u^*) - u| > \frac{1}{n}\} \leq 6n^2\epsilon_n u.$$

Since $\sum p_n$ is finite, the desired result follows from the Borel-Cantelli lemma.

As Chung comments in [3], the preceding proof is in the grand tradition of classical probability. But then, what of the result?

3. The moments of excursionary occupation

In this section we use the results in Chung [2], beginning with a review of the notation. Let

$$\gamma(t) = \sup\{s \mid s \le t; B(s) = 0\}$$

$$\beta(t) = \inf\{s \mid s \ge t; B(s) = 0\}$$

$$\lambda(t) = \beta(t) - \gamma(t).$$

Then $(\gamma(t), \beta(t))$ is the excursion interval straddling t, and $\lambda(t)$ is its duration. For any Borel set A in $[0, \infty)$:

$$S(t; A) = \int_{\gamma(t)}^{\beta(t)} 1_A(|B(u)|)\, du$$

is the occupation time of A by $|B|$ during the said excursion. Its expectation conditioned on $\gamma(t)$ and $\lambda(t)$ has a density given by

$$\mathcal{E}(S(t; dx) \mid \gamma(t) = s, \lambda(t) = a) = 4xe^{-2x^2/a}dx. \tag{3.1}$$

This result is due to Lévy; a proof is given in [2]. Integration gives

$$\mathcal{E}(S(t; (0, \epsilon)) \mid \gamma(t) = s, \lambda(t) = a) = a(1 - e^{-2\epsilon^2/a}). \tag{3.2}$$

Next it follows from (2.22) and (2.23) in [2] that

$$P\{\lambda(t) \in da\} = \frac{1}{\pi}\sqrt{t/a^3}\, da \quad \text{for } a \ge t. \tag{3.3}$$

In particular if $r > c \ge t > 0$ then $P\{\lambda(t) > c\} > 0$ and

$$P(\lambda(t) \in dr \mid \lambda(t) > c) = \frac{1}{2}\sqrt{c/r^3}\, dr. \tag{3.4}$$

Lévy derived (3.4) from the property of the Lévy process T described in section 1 above. It is a pleasure to secure this fundamental result directly from our excursion theory.

What is the exact relation between the excursion straddling t and the sequence of excursions (φ_n) introduced in Section 2?

Recall that φ_n is the nth excursion of duration exceeding c for given $c > 0$. We claim that φ_1 is distributed like the excursion straddling c conditional on its duration exceeding c. To see this we introduce a new sequence of excursions (η_n) with excursion intervals (γ_n, β_n) of duration $\lambda_n = \beta_n - \gamma_n$. Define η_1 as the excursion straddling $t = c$ with excursion interval (γ_1, β_1); then define η_2 as the excursion straddling $t = \beta_1 + c$ with excursion interval (γ_2, β_2); η_3 as the excursion straddling $t = \beta_2 + c$, etc. Note that the post-β_1 process $B_1(t) = B(\beta_1 + t)$, is a Brownian motion which is independent of the excursion η_1. As in Theorem 2.1 a simple induction argument shows that the sequence (η_n) is i.i.d., at least if we shift the excursions so as to start at $t = 0$. Since for any sample point ω in the underlying probability space $\varphi_1(\omega)$ is the first element of the sequence $(\eta_n(\omega))$ of duration exceeding c, it follows that φ_1 is distributed like the excursion straddling c, conditional on its duration exceeding c.

Now we can compute by (3.2) and (3.4):

$$\frac{1}{\sqrt{c}}\mathcal{E}(S(t;(0,\epsilon)) \mid \lambda(t) > c) = \int_c^\infty r(1 - e^{-2\epsilon^2/r})\frac{dr}{2r^{3/2}}$$

$$\to \int_0^\infty r(1 - e^{-2\epsilon^2/r})\frac{dr}{2r^{3/2}} = \epsilon\sqrt{2\pi} \qquad \text{as } c \to 0.$$

This is (2.1) if we choose $t = c$.

Next Chung proved as a particular case of Theorem 9 in [2]:

$$\mathcal{E}(S(t;(0,\epsilon))^k \mid \gamma(t) = s, \lambda(t) = a) \le (k+1)!\,\epsilon^{2k} \qquad k \ge 1. \tag{3.5}$$

For $k = 2$ this is the missing estimate mentioned in Section 0. But it is also trivial that

$$S(t;(0,\epsilon)) \le \lambda(t). \tag{3.6}$$

Using (3.4), (3.5) and (3.6) we have

$$\begin{aligned}
\mathcal{E}(S(t;(0,\epsilon))^2 \mid \lambda(t) > c) &= \int_c^\infty \mathcal{E}(S(t;(0,\epsilon))^2 \mid \lambda(t) = r)\frac{\sqrt{c}\,dr}{2r^{3/2}} \\
&\le \int_0^\infty (6\epsilon^4 \wedge r^2)\frac{\sqrt{c}\,dr}{2r^{3/2}} \\
&\le \sqrt{c}\left(6\epsilon^4 \int_{4\epsilon^2}^\infty \frac{dr}{2r^{3/2}} + \int_0^{4\epsilon^2} \frac{\sqrt{r}}{2}dr\right) \\
&\le 6\epsilon^3\sqrt{c}.
\end{aligned}$$

Now choose $t = c$. Then $S(c;(0,\epsilon))$ conditional on $\lambda(c) > c$ is distributed like $f_\epsilon(|\varphi_1|)$. Hence

$$\mathcal{E}(X_n^2) \le 6\sqrt{c}\epsilon^3/\epsilon^2 = 6\epsilon\sqrt{c}.$$

This is (2.2).

References

[1] P. LÉVY, Sur certains processus stochastiques homogènes. *Compositio Math.* 7 (1939), 283-339.

[2] K.L. CHUNG, Excursions in Brownian motion. *Arkiv för Mat.* 14 (1976), 155-177.

[3] K.L. CHUNG, Reminiscences of some of Paul Lévy's Ideas in Brownian Motion and in Markov Chains. *Seminar on Stochastic Processes 1988*, 99-108. Birkhäuser, 1989. Also printed with the author's permission but without the Postscript in *Colloque Paul Lévy*, Soc. Math. de France, 1988.

[4] C. DELLACHERIE & P.A. MEYER, *Probabilités et Potentiel.* Chapitres XII à XVI, Hermann, Paris, 1987.

[5] P. LÉVY, *Processus Stochastiques et Mouvement Brownien.* Gauthier-Villars, Paris, 1948 (second edition 1965).

[6] J.F.C. KINGMAN, *Regenerative Phenomena.* Wiley, New York, 1972.

[7] P. GREENWOOD & J. PITMAN, Construction of local time and Poisson point processes from nested arrays. *J. London Math. Soc.* (2) 22 (1980), 182-192.

A.A. Balkema
F.W.I., Universiteit van Amsterdam
Plantage Muidergracht 24
1018 TV Amsterdam, Holland

K.L. Chung
Department of Mathematics
Stanford University
Stanford, CA 94305

Potential Analysis 1: 83–92, 1992.

Greenian Bounds for Markov Processes

KAI LAI CHUNG
Department of Mathematics, Stanford University, Stanford, California 94305-2125, U.S.A.

(Received: 3 July 1991; accepted: 14 August 1991)

Abstract. In Section 1, a temporal bound is estimated by a spatial bound, for a Markov process whose transition density satisfies a simple condition. This includes the Brownian motion, for which comparison with a more special method is made. In Section 2, the result is related to the Green operator and examples are given. In Section 3, the result is applied to an old problem of eigenvalues of the Laplacian. In Section 4, recent extensions from the Laplacian to the Schrödinger circle-of-ideas are briefly described. In this case, time is measured by an exponential functional of the process, commonly known under the names Feynman–Kac.

This article was written as script for two talks, one general and one technical, given in Taiwan in January 1991. In the spirit of the occasion, the style of exposition is deliberately paced and discursive.

AMS subject classifications (1991). Primary: 60J45, 60J36; Secondary 60J40, 35J10.

Key words. Markov process, exit time, Green operator, Laplacian.

Section 1

Let $\{X_t, t \geqslant 0\}$ be a Markov process with the filtration $\{\mathscr{F}_t, t \geqslant 0\}$ and transition semigroup $\{P_t, t \geqslant 0\}$. Let S denote its state space, $\mathscr{B}(S)$ the Borel tribe. For $D \in \mathscr{B}(S)$, define the random variable

$$\tau_D = \inf\{t > 0 \mid X_t \in S \backslash D\}.$$

This is called the 'exit time' of D. Under general hypotheses τ_D is 'optional' in the sense that for each $t > 0$:

$$\{\tau_D > t\} \in \mathscr{F}_{t+}.$$

This is a fundamental result in the theory of Hunt processes, and it is not easy to prove! If we restrict ourselves to a closed set D, the result is trivial; but even for an open set D there is some difficulty. We shall assume it here, together with the Markov property of $\{X_t\}$ with respect to $\{\mathscr{F}_{t+}\}$. The latter is contained in the usual form of the strong Markov property, valid for Hunt processes.

Now we make a special asumption about the transition probability P_t. Let m be a σ-finite measure on $\mathscr{B}(S)$. We assume that for each $t > 0$, P_t has a density $p_t(\cdot,\cdot)$: so that for each $x \in S$ and $B \in \mathscr{B}(S)$ we have

$$P_t(x, B) = \int_B p_t(x, y) m(\mathrm{d}y).$$

Furthermore, we assume that there are two constants A and α, $0 < A < \infty$, $0 < \alpha < \infty$, such that for each $t > 0$ we have for all x and y:

$$p_t(x, y) \leqslant \frac{A}{t^\alpha}. \tag{1}$$

These asumptions are satisfied when X is the Brownian motion process in R^d ($d \geqslant 1$). In this case m is the Lebesgue measure, and

$$p_t(x, y) = \frac{1}{(2\pi t)^{d/2}} e^{-\frac{\|x-y\|^2}{2t}};$$

so that we may take in (1)

$$A = \frac{1}{(2\pi)^{d/2}}, \quad \alpha = \frac{d}{2}.$$

In fact, the general framework set up here is motivated by this particular case. There is no reason not to make the generalization since it even adds to the clarity of the calculations below. It will also become clear that the condition (1) is of a *genre* that is easy to tinker with, tant mieux!

We proceed to the probabilistic estimation of τ_D. For each $x \in S$ and $t > 0$, we have obviously:

$$\begin{aligned} P^x\{\tau_D > t\} \leqslant P^x\{X_t \in D\} &= \int_D p_t(x, y) m(\mathrm{d}y) \\ &\leqslant \int_D \frac{A}{t^\alpha} m(\mathrm{d}y) = \frac{Am(D)}{t^\alpha}. \end{aligned} \tag{2}$$

Call the last numbr θ and choose t preliminarily to make $\theta < 1$:

$$\theta = \frac{Am(D)}{t^\alpha} < 1. \tag{3}$$

Thus

$$\sup_x P^x\{\tau_D > t\} \leqslant \theta,$$

where the sup is over S. The next step is the crucial application of the Markov property. For each integer $n \geqslant 1$, we have[1]

$$
\begin{aligned}
P^x\{\tau_D > nt\} &= P^x\{\tau_D > (n-1)t; P^{X_{(n-1)t}}[\tau_D > t]\} \\
&\leqslant P^x\{\tau_D > (n-1)t\} \sup_x P^x\{\tau_D > t\} \leqslant P^x\{\tau_D > (n-1)t\}\theta.
\end{aligned}
$$

Observe that in the equation above, we have applied the Markov property at time $(n-1)t$, with respect to $\mathscr{F}_{(n-1)t+}$ (and have written P^x for the older notation E^x). It follows by induction on n that

$$
\sup_x P^x\{\tau_D > nt\} \leqslant \theta^n. \tag{4}
$$

A neat way of obtaining an inequality for $E^x\{\tau_D\}$ is shown below:

$$
E^x\left\{\frac{\tau_D}{t}\right\} \leqslant \sum_{n=0}^{\infty} P^x\left\{\frac{\tau_D}{t} > n\right\} \leqslant \sum_{n=0}^{\infty} \theta^n = \frac{1}{1-\theta}.
$$

Thus

$$
\sup_x E^x\{\tau_D\} \leqslant \frac{t}{1-\theta} = \frac{t}{1 - Am(D)t^{-\alpha}}, \tag{5}
$$

and it remains to choose t as best we can. This is easy to do by elementary calculus. If we denote by $\phi(t)$ the function of t in the last member of (5), we find that $\phi'(t_0) = 0$ where $t_0^\alpha = (1+\alpha)Am(D)$, and the corresponding $\theta_0 = 1/(1+\alpha)$. Thus ϕ attains its minimum in $((Am(D))^{1/\alpha}, \infty)$ at t_0 and

$$
\phi(t_0) = (1+\alpha)^{1+(1/\alpha)} \frac{1}{\alpha} A^{1/\alpha} m(D)^{1/\alpha}.
$$

Therefore we have proved the following result.

THEOREM 1. $\sup_x E^x\{\tau_D\} \leqslant (1+\alpha)^{1+(1/\alpha)} \dfrac{1}{\alpha} A^{1/\alpha} m(D)^{1/\alpha}.$ (6)

In the case of the Brownian motion process in R^d, the righthand member in (6) becomes

$$
\frac{1}{2\pi}\frac{d+2}{d}\left(\frac{d+2}{2}\right)^{2/d} m(D)^{2/d}. \tag{7}
$$

Now if D is a ball $B(0,r)$ with center at the origin and radius r, a well-known argument via martingale theory yields the exact formula

$$
E^x\{\tau_{B(0,r)}\} = \frac{r^2 - \|x\|^2}{d}
$$

for $x \in B(0,r)$; while the lefthand member is equal to zero for $x \notin B(0,r)$. Hence we have

$$\sup_x E^x\{\tau_{B(0,r)}\} = \frac{r^2}{d}.$$

We know that

$$m(B(0,r)) = \frac{2\pi^{d/2}}{d\Gamma\left(\frac{d}{2}\right)} r^d.$$

Substituting into the above we obtain

$$\sup_x E^x\{\tau_{B(0,r)}\} = \frac{1}{\pi d}\left(\frac{d}{2}\right)^{2/d} \Gamma\left(\frac{d}{2}\right)^{2/d} m(B(0,r))^{2/d}. \tag{8}$$

For $d = 2$ the righthand member of (8) reduces to $(1/2\pi)m(B(0,r))$, while that in (7) reduces to $(2/\pi)m(D)$. Thus the general estimate is 4 times larger. An interesting computation (which is recommended to the reader) shows that as $d \to \infty$, the general estimate becomes e^2 times larger than the exact value for the ball. It is known that among sets D in R^d of a given volume, the quantity $\sup_x E^x\{\tau_D\}$ attains its maximum when D is a ball.

Section 2

Next, we proceed to put the matter in another context. We begin by considering the integral

$$\int_0^{\tau_D} f(X_t)\,dt.$$

If $t \to X_t$ is Borel measurable, and $f \in \mathscr{B}(S)$, $f \geq 0$, then the integral is defined for each ω (hidden in the notation!) in the Lebesgue sense, but may be $+\infty$. Moreover we can define its mathematical expectation as follows:

$$G_D f(x) = E^x\left\{\int_0^{\tau_D} f(X_t)\,dt\right\}. \tag{9}$$

We call this the Green operator for D. This is defined when X is a Hunt process because then $t \to X_t$ is right-continuous, hence Borel measurable. It is customary to assume that D is an open set, or even an open and connected set. It is clear that we may replace f in (9) by $1_D f$ where 1_D is the indicator of D, so that f may be defined only on D; but it is often a moot question whether we wish to restrict the x in (9) also to D. Our first observation is that

$$G_D 1(x) = E^x\{\tau_D\};$$

and consequently

$$||G_D 1|| = \sup_x E^x\{\tau_D\} \tag{10}$$

where $||\ldots||$ denotes the sup-norm. Incidently, it is not hard to show that the sup in (10) may be taken over S or over D, indifferently. Thus the moot question mentioned above causes no problem here.

After (9) is defined for $f \geqslant 0$, we can extend it to a general f by the standard procedure: $f = f^+ - f^-$, etc. Thus $G_D f(x)$ is defined and finite if and only if $G_D|f|(x)$ is finite, just as in the case of a Lebesgue integral. It follows that if $||G_D 1|| < \infty$ then $||G_D f|| < \infty$ for any bounded Borel measurable f. It turns out that the condition $||G_D 1|| < \infty$ is of sufficient importance to deserve a name.

DEFINITION. The set D in $\mathscr{B}(S)$ is called Green-bounded in case $||G_D 1|| < \infty$.

Thus Theorem 1 asserts that if $m(D) < \infty$ then D is Green-bounded. We begin with a few comments on the definition above.

(i) In R^1, the only open and connected set which is Green-bounded is a bounded open interval.

Proof. The only open and connected sets D in R^1 with $m(D) < \infty$ are bounded open intervals. Such a set is Green-bounded by Theorem 1. The open and connected sets with $m(D) = \infty$ are of the form $(-\infty, a)$ or (a, ∞) where $-\infty < a < \infty$, or $R^1 = (-\infty, +\infty)$ itself. Consider $D = (a, \infty)$; then under P^x where $x \in (a, \infty)$ we have

$$\tau_D = T_a = \inf\{t > 0 | X_t = a\}.$$

Thus for any $x > a$:

$$G_D 1(x) = E^x\{T_a\} = +\infty$$

by a well-known (but astonishing) result in Brownian motion on the line. Alternatively the 'Green Function' (namely the density kernel of the Green operator) of (a, ∞) is well known:

$$G_{(a,\infty)}(x, y) = 2(x - a) \wedge 2(y - a), \quad x \in (a, \infty), \quad y \in (a, \infty).$$

Hence for each $x > a$:

$$\int_a^\infty G_{(a,\infty)}(x, y)\,\mathrm{d}y = \int_a^\infty [2(x - a) \wedge 2(y - a)]\,\mathrm{d}y$$

$$= \int_a^x (y - a)\,\mathrm{d}y + (x - a)\int_x^\infty \mathrm{d}y = +\infty.$$

Similarly for $D=(-\infty,a)$. As for $D=(-\infty,\infty)$, we have $\tau_D=+\infty$ and so $E^x\{\tau_D\}=+\infty$. Therefore we have shown that an open and connected set D in R^1 with $m(D)=\infty$ is not Green-bounded.

(ii) There is a Green-bounded, open and connected set D in R^d, $d\geqslant 2$, such that $m(D)=\infty$.

Proof. In R^2, consider

$$D=\{(x',x'')\,|\,a<x'<b;\ -\infty<x''<\infty\}$$

where $-\infty<a<b<+\infty$. To see that D is Green-bounded, we recall that the two components of $X=(X',X'')$ are independent Brownian motion processes in R^1. It follows (why?) that

$$E^{(x',x'')}\{\tau_D\}=E^{x'}\{\tau_{(a,b)}\}$$

which is bounded for all $x=(x',x'')$, by Theorem 1 since $m((a,b))<\infty$.

(iii) There is a negative property of Green-boundedness which is worthy of notice. A characteristic property of a set D with $m(D)<\infty$ is that for any $\varepsilon>0$ there exists a compact subset C of D such that $m(D\backslash C)<\varepsilon$. But it is false if $\|G_D1\|<\infty$ then such a subset exists with $\|G_{D\backslash C}1\|<\varepsilon$. The set D given under (ii) is an example. Indeed for any compact subset C of D, we have

$$\|G_{D\backslash C}1\|=\|G_D1\|.$$

To see this, observe that $D\backslash C$ contains the set below:

$$B=\{(x',x'')|a'<x'<b;\ x''>c\}$$

for all sufficiently large c, and it is clear that

$$E^{(x,x'')}\{\tau_{D\backslash C}\}\geqslant E^{(x',x'')}\{\tau_B\}=E^{x'}\{\tau_{(a,b)}\}=E^{(x',x'')}\{\tau_D\}.$$

Hence the '$\geqslant$' above is actually '$=$'.

The property of Green-boundedness of a set is a meaningful extension of the property of its having a finite measure, in certain questions concerning the 'size' of the set. This is not surprising because for a stochastic process it is often the time element that plays the essential role. Indeed, the rndom time τ_D is called the 'life-time' of 'the process X killed outside the set D'. Such a time is never bounded in the literal sense, namely for (almost) all ω, except possibly in trivial cases, but when it is bounded in the Green sense as defined above, it possesses remarkable properties. Some recent results of this kind are given in [3].

Section 3

Here we will relate the Green bound to a popular object in classical and applicable analysis. For simplicity's sake we confine the discussion to the case where X is the Brownian motion process.

Consider a bounded and smooth domain D in $R^d (d \geqslant 1)$. A domain is an open and connected set; it is 'smooth' if its boundary ∂D satisfies some differentiability conditions so that certain general transformation formulas of integral calculus are applicable; see below for a specification. More generally, ∂D is 'regular' in the probabilistic sense when $E^z\{\tau_D\} = 0$ for all $z \in \partial D$. But we shall not pursue this generality rarely treated in analysis. Indeed, a non-probabilitistic definition of regularity would take several paragraphs to write down.

Let Δ denote the Laplacian (operator) in R^d:

$$\Delta = \sum_{i=1}^{d} \left(\frac{\partial}{\partial x_i}\right)^2.$$

The real number λ is called an eigenvalue for D and Δ iff there exists a function $\phi \not\equiv 0$ such that

$$\begin{aligned} \Delta\phi &= \lambda\phi \text{ in } D; \\ \phi &= 0 \text{ on } \partial D. \end{aligned} \tag{11}$$

If so, then ϕ is called an eigenfunction corresponding to λ. We will treat the equation (11) in the strict sense, namely $\phi \in C_0^2(D)$, where $C_0^2(D)$ denotes the class of f on D such that f is twice continuously differentiable in D and $f(x) \to 0$ as $x \in D$, $x \to \partial D$. Under the conditions in (11), we can prove that

$$G_D(\Delta\phi) = -2\phi \tag{12}$$

in D, without even assuming any regularity of D. Such a proof has been communicated to me by Ruth Williams using stochastic integration. Observe that we are not assuming the Hölder continuity of $\Delta\phi$ and so the result is not "obvious", and is a good exercise.

Using (12) we can transform (11) into

$$-2\phi = G_D(\Delta\phi) = G_D(\lambda\phi)$$

or

$$\lambda G_D\phi = -2\phi. \tag{13}$$

Let us first prove that $\lambda < 0$.[3] If ∂D is smooth enough to allow the use of Green's (first) formula in calculus, then we have by (11), in customary notation:

$$\int_D \lambda\phi^2 = \int_D \phi\Delta\phi = -\int_D (\nabla\phi)^2 + \int_{\partial D} \phi\frac{\partial\phi}{\partial n} = -\int_D (\nabla\phi)^2$$

since $\phi = 0$ on ∂D. Hence $\lambda < 0$ since $\phi \not\equiv \forall$ in D. It follows from (13) that for any x in D we have

$$-\lambda = \frac{2\phi(x)}{G_D\phi(x)} = \frac{2|\phi(x)|}{|G_D\phi(x)|}. \tag{14}$$

Since $|\phi| \in C_0(D)$, there exists $x_0 \in D$ such that

$$|\phi(x_0)| = \|\phi\|.$$

Consequently putting $x = x_0$ in (14) we obtain

$$-\lambda = \frac{2\|\phi\|}{|G_D\phi(x_0)|} \geqslant \frac{2\|\phi\|}{\|G_D 1\| \, \|\phi\|} = \frac{2}{\|G_D 1\|}$$

or

$$\lambda \leqslant -\frac{2}{\|G_D 1\|}. \tag{15}$$

Observe that we did not use the known result that $\phi > 0$ in D, partly because we are dubious about its usual proof given in textbooks. Thus an upper bound for the $\|G_D 1\|$ in Theorem 1 yields an upper bound for all possible eigenvalues. In particular if $m(D) \to 0$ it follows that the largest eigenvalue must go to $-\infty$.

Let us note that the preceding argument is independent of any known theory of the so-called 'spectrum' of the Laplacian. Whereas the latter is an ancient and powerful tool of analysis, those who tend to be preoccupied with it would do well to study the newer approaches offered by probability theory as illustrated here.

Section 4

We concloude with a brief sketch of a recent extension of the preceding consideration. For simplicity's sake let X be the Brownian motion process in what follows.

Let J denote the class of functions q in $\mathscr{B}(S)$ satisfying the following condition:

$$\lim_{t \downarrow 0} \sup_{x \in S} E^x \left\{ \int_0^t |q|(X_s) \, ds \right\} = 0;$$

and put for $t \geqslant 0$:

$$e_q(t) = e^{\int_0^t q(X_s) \, ds}.$$

Using $e_q(\cdot)$ as the multiplicative functional we define a semigroup $\{T_t, t \geqslant 0\}$ by

$$T_t f(x) = E^x\{e_q(t) f(X_t)\}$$

where f is the usual symbol. See [2] for some information regarding $\{T_t\}$. We can prove that there exist two constants $C_0 > 0$ and $C_1 > 0$ such that for all $t \geqslant 0$:

$$\|T_1 1\| \leqslant e^{C_0 + C_1 t}.$$

Now the Green operator in (9) is generalized to

$$G_D^{(q)} f(x) = E^x\left\{\int_0^{\tau_D} e_q(t) f(X_t)\,dt\right\}$$

which may be called the q-Green operator. There is an important difference: even when D is bounded $G_D^{(q)}1$ may be identically $+\infty$ in D, contrary to the case when $q \equiv 0$. But the following result is true.

If $m(D) < \infty$,[2] and

$$G_D^{(q)} 1(x) = E^x\left\{\int_0^{\tau_D} e_q(t)\,dt\right\}$$

is finite for at least one x in D, then $\|G_D^{(q)}1\| < \infty$: namely D is 'q-Green bounded'. In this case the 'gauge' function for (D, q):

$$u(x) = E^x\{e_q(\tau_D)\}$$

is also bounded (in R^d). Why should this be interesting? Because when D is a regular bounded domain, and its gauge is bounded, then the Dirichlet boundary value problem for the Schrödinger equation:

$$\left(\frac{\Delta}{2} + q\right)\phi = 0$$

in D has a unique solution (in the 'weak' sense) for any given continuous boundary function f on ∂D. Last but most important, the solution is given by the explicit probabilistic formula:

$$u_f(x) = E^x\{e_q(\tau_D) f(X(\tau_D))\},\ x \in D.$$

This representation is most expedient in various applications; see [4].

Obviously this is a long story which cannot be told here. For the particular case when q is bounded, see the last part of Chapter 4 in [1]. For the general case, except for the connection between $G_D^{(q)}1$ and u above, see [3] for further information and references.

Notes

[1] Replacing t by $t - 1|/n$ in (2) and letting $n \to \infty$ we obtain an improvement of (2) where $P^x\{\tau_D > t\}$ is replaced by $P^x\{\tau_D \geqslant t\}$.
[2] The necessity of this stronger condition than $\|G_D 1\| < \infty$ is due to the comment (iii) in Section 2.
[3] A better proof of this was communicated to me by Ruth Williams and V. Papanicolaou. Suppose $\lambda \geqslant 0$, then by the maximum principle for elliptic PDE, if $\phi \in C^2(D) \cap C(\bar{D})$, $\phi = 0$ on ∂D, $(\Delta - \lambda)\phi = 0$ in D, then $\phi = 0$ in $\bar{D}$. No regularity of D is needed.

References

1. Lectures From Markov Processes to Brownian Motion, *Grundlehren der Mathematischen Wissenschaften* **249**, Springer-Verlag, 1982.
2. Doubly-Feller Process with Multiplicative Functional, *Seminar on Stochastic Processes 1985*, pp. 63–78, Birkhäuser.
3. Gauge Theorem for Unbounded Domains, *Seminar on Stochastic Processes 1988*, 87–98, Birkhäuser.
4. Probability Methods in Potential Theory, Potential Theory Surveys and Problems, Proceedings, Prague 1987. *Lecture Notes in Mathematics* **1344**, 42–54, Springer-Verlag, 1987.

Journal of Theoretical Probability, Vol. 6, No. 3, 1993
With kind permission of Springer Science and Business Media.

Doeblin's Big Limit Theorem

Kai Lai Chung[1,2]

Received January 15, 1992

Doeblin regarded his paper[2] as his hardest work. The big limit is that of $P^{(n)}(x, E)$ as n tends to infinity, in a measurable non-topologized space. An exposition of part one of this paper was published in Ref. 1. This is the exposition of part two, which contains some reparation as well as clarification.

KEY WORDS: Markov processes; limit theory.

1. INTRODUCTION

This is the second part of an exposition of Ref. 2. The first part has appeared in Ref. 1. The manuscript for the second part was dated May 31, 1964. It could have been published together with the first part, but I withheld it in the hope of making further improvements in the exposition, after giving a course on both parts in Stanford University around that time. An earlier version was given in my lectures in Columbia University, in the spring semester of 1951. [The late Professor Abraham Wald would have been in the audience had he not died in an airplane accident in India shortly before my course began.] The notes were dittoed and made available to the public. Let me quote some of the appended remarks, slightly altered.

This remarkable work, in which Doeblin set forth his most general theory of (discrete time) Markov processes, contains a wealth of ideas and methods, and is, to our mind, unsurpassed in its depth of probabilistic analysis. His style, however, is sometimes negligent. We have undertaken to sift out his results and give them the shortest complete proofs known to us.

[1] Department of Mathematics, Stanford University, Stanford, California 94305-2125.

[2] Doeblin's great contributions to probability theory were remembered at a conference "50 Years after Doeblin: Developments in the Theory of Markov Chains, Markov Processes and Sums of Random Variables" held at Blaubeuren, Germany, November 2-7, 1991. Professor Chung contributed this article to further celebrate this occasion. The editor thanks him.

0894-9840/93/0700-0417$07.00/0

Several gaps in the paper were filled in at considerable pains. First, the whole treatment of "cycles" (see Definition 13 in Ref. 1) has been clarified and strengthened in Ref. 1. Next, in the proof of Proposition 53, the fact that the sequence $(n_{i,b})$ depends on b seems to have been disregarded by Doeblin (cf. p. 95, Ref. 2), which necessitated our more elaborate construction there. In this and some other instances, it is impossible to tell whether Doeblin was unaware of the gaps or whether he had actually short-cuts which he omitted.

Let me also quote from Doeblin himself, from his letter to Fréchet dated October 28, 1939:

> "(which) is certainly after the general theory of chains the most difficult problem I have been able to solve."

"Which" refers to his work on "the set of powers" of a given probability law, and the literal translation is mine.

As mentioned in my old notes, there is a third of Ref. 2 which deals with the case when the hypothesis (H) in Proposition 56 is not satisfied, and which he called the "anormal" case. In resurrecting the present manuscript, I found to my amazement a stack of time-yellowed handwritten long sheets which apparently contained an exposition of this last part of Ref. 2. But it must be regarded as "inedit," and anyone who is interested will be better advised to try Doeblin's own account.

Postscript. B. Bru has informed me that he found in the archives of Marbach a letter from Doeblin to S. Polig in which he said:

> "Darunter befindet sich meine beste Arbeit...",

referring to his manuscript of Ref. 2 here. Moreover, he said, in my translation from the German:

> "whereas I believe I can recover most of my other stuff in case of destruction of the manuscript, I don't know if I should be able to do so easily with this work. Therefore I should like to have a draft of my manuscript put in relative security."

In what follows, we continue the Doeblin's theory as presented in Ref. 1. The symbols, notations and definitions of various terms/concepts used here are as in Ref. 1, and are continued without any further explanation.

X is indecomp. and abs. ess. We write $A \in \mathscr{A}$ for "A is abs. ess."

Proposition 51. Let $E \in \mathscr{B}$ be given and suppose that $\exists A \in \mathscr{A}$ and a real number λ such that

$$\forall x \in A: \overline{\lim_{n}}\, P^{(n)}(x, E) \leqslant \lambda \tag{1}$$

Then (1) holds for every $x \in X - F$ where $F \notin \mathscr{A}$.

Proof. By Proposition 19, $A^\infty \neq \varnothing$ and by Proposition 18, $X - A^\infty \notin \mathscr{A}$. Let $x \in A^\infty$, then $\forall m \geqslant 1$, if $n > m$:

$$P^{(n)}(x, E) \leqslant \sum_{j=1}^{m} \int_A P^{(n-j)}(y, E)\, K^{(j)}(x, dy) + \sum_{j=m+1}^{m} K^{(j)}(x, A)$$

If follows by Fatou's lemma that

$$\begin{aligned}\overline{\lim_n}\, P^{(n)}(x, E) &\leqslant \sum_{j=1}^{m} \overline{\lim_n}\, P^{(n-j)}(y, E)\, K^{(j)}(x, dy) + \sum_{j=m+1}^{\infty} K^{(j)}(x, A) \\ &\leqslant \lambda \sum_{j=1}^{m} K^{(j)}(x, A) + \sum_{j=m+1}^{\infty} K^{(j)}(x, A)\end{aligned}$$

Since $L(x, A) = \sum_{j=1}^{\infty} K^{(j)}(x, A) = 1$, this implies (1) if we take $F = X - A^\infty$.

Proposition 52. For each $F \in \mathscr{B}$, there exists a set $F(E) \notin \mathscr{A}$ such that $\forall x \notin F(E)$, the upper-limit $\overline{\lim}_n\, P^{(n)}(x, E)$ is the same number $\bar{P}(E)$.

Proof. For each λ, $0 \leqslant \lambda \leqslant 1$, we put

$$E_\lambda = \{x \in X \mid \overline{\lim_n}\, P^{(n)}(x, E) \leqslant \lambda\}$$

and

$$\bar{\lambda} = \inf\{\lambda \mid E_\lambda \in \mathscr{A}\}$$

Thus for each $m \geqslant 1$, we have

$$E_{\bar{\lambda}+m^{-1}} - E_{\bar{\lambda}-m^{-1}} \in \mathscr{A}$$

Here by Proposition 51, $\exists F_m \notin \mathscr{A}$ such that

$$\forall x \notin F_m\colon \bar{\lambda} - \frac{1}{m} < \overline{\lim_n}\, P^{(n)}(x, E) \leqslant \bar{\lambda} + \frac{1}{m}$$

Let $C_m = F_m^c \cap F_m^0$, then C_m is cl. by Proposition 19 and Proposition 1. Let $C = \bigcap_{m=1}^{\infty} C_m$, then C is cl. and $X - C \notin \mathscr{A}$ by Proposition 18. We have

$$\forall x \in C\colon \overline{\lim_n}\, P^{(n)}(x, E) = \bar{\lambda} \tag{2}$$

Notation. For each $E \in \mathscr{B}$, let the set C in Eq. (2) be denoted by $C(E)$ and let the number $\bar{\lambda}$ be denoted by $\bar{P}(E)$.

Definition. For a fixed positive integer D and a sequence of integers $\{n_i\} \uparrow \infty$ let $\rho(\{n_i\})$ be the greatest integer ρ such that $n_i, n_i+D, \ldots, n_i+(\rho-1)D$ are all members of the sequence for infinitely many values of i; if no such ρ exists, we set $\rho(\{n_i\}) = +\infty$. For each $E \in B$ and $x \in C(E)$, we define

$$\rho_x = \rho_x(D, E) = \sup\{\rho(\{n_i\}) \mid \overline{\lim_n} P^{(n_i)}(x, E) = \bar{P}(E)\}$$

Proposition 53. Let $x \in C(E)$, then there exists a set $F = F(x, E) \notin \mathscr{A}$ such that

$$\forall y \notin F(x, E): \rho_y \geqslant \rho_x$$

Proof. By Proposition 52, $\exists\{n_i\} \uparrow$ such that

$$\lim_i P^{(n_i)}(x, E) = \bar{P}(E)$$

Put for each $m \geqslant 1$:

$$F_{i,m} = \left\{ y \in C(E) \,\middle|\, P^{(n)}(y, E) > \bar{P}(E) + \frac{1}{4^i} \text{ for some } n \geqslant m \right\}$$

For each i, we have $\bigcap_{m=1}^{\infty} F_{i,m} = \varnothing$. For every integer $b \geqslant 1$, $\exists$ a subsequence $\{n_{i,b}, i \geqslant 1\}$ of $\{n_i\}$ satisfying:

(i) $P^{(b)}(x, F_{i,n_{i,b}-b}) < 1/4^i$;

(ii) $P^{(n_{i,b})}(x, E) > \bar{P}(E) - 1/4^i$;

(iii) $\rho(\{n_{i,b}\}) = \rho_x$;

(iv) $\{n_{i,b}\}$ is a subsequence of $\{n_{i,b-1}\}$ where $\{n_{i,0}\} = \{n_i\}$.

This is easily seen by induction on b. Having chosen $\{n_{i,b}\}$, we put

$$G_{i,b} = \left\{ y \in C(E) \,\middle|\, P^{(n_{i,b}-b)}(y, E) < \bar{P}(E) - \frac{1}{2^i} \right\}$$

We have then

$$\begin{aligned}
\bar{P}(E) - \frac{1}{4^i} &< P^{(n_i,n)}(x, E) \\
&= \left(\int_{G_{i,b}} + \int_{G_{i,b}^c \cap F_{i,n_{i,b}-b}} + \int_{G_{i,b}^c \cap F_{i,n_{i,b}-b}^c} \right) P^{(n_{i,b}-b)}(y, E)\, P^{(b)}(x, dy) \\
&\leqslant P^{(b)}(x, G_{i,b}) \left(\bar{P}(E) - \frac{1}{2^i} \right) + \frac{1}{4^i} + (1 - P^{(b)}(x, G_{i,b})) \left(\bar{P}(E) + \frac{1}{4^i} \right) \\
&\leqslant \bar{P}(E) + \frac{2}{4^i} - \frac{1}{2^i} P^{(b)}(x, G_{i,b})
\end{aligned}$$

Thus

$$P^{(b)}(x, G_{i,b}) \leqslant \frac{3 \cdot 2^i}{4^i} = \frac{3}{2^i}$$

Put for $l \geqslant 1$:

$$R_{l,b} = \bigcup_{i=l}^{\infty} G_{i,b}, \qquad R_b = \bigcap_{l=1}^{\infty} R_{l,b}, \qquad R = \bigcap_{b=1}^{\infty} R_b$$

We have by definition ($C = C(E)$):

$$\forall y \in C - R_{l,b}: \qquad P^{(n_{i,b}-b)}(y, E) \geqslant \bar{P}(E) - \frac{1}{2^i} \qquad \text{for all} \quad i \geqslant l \tag{3}$$

Since

$$P^{(b)}(x, R_{l,b}) \leqslant \sum_{i=l}^{\infty} \frac{3}{2^i} = \frac{3}{2^{l-1}}$$

we have

$$P^{(b)}(x, R_b) = 0$$

and

$$\forall b \geqslant 1: P^{(b)}(x, R) = 0$$

Hence $L(x, R) = 0$, $R^0 \neq \varnothing$, $R \notin \mathscr{A}$ by Proposition 19. If $y \in C - R$, then $\exists\, l$ and b:

$$P^{(n_{i,b}-b)}(y, E) \geqslant \bar{P}(E) - \frac{1}{2^i} \qquad \text{for all} \quad i \geqslant l$$

Hence

$$\varliminf_i P^{(n_{i,b}-b)}(y, E) \geqslant \bar{P}(E) \geqslant \varlimsup_i P^{(n_{i,b}-b)}(y, E) \tag{4}$$

and consequently

$$\forall y \in C - R, \exists b: \lim_i P^{(n_{i,b}-b)}(y, E) = \bar{P}(E)$$

Furthermore by (iii),

$$\rho_y \geqslant \rho(\{n_{i,b} - b\}) = \rho(\{n_{i,b}\}) = \rho_x$$

Thus, Proposition 53 is proved if we take $F(x, E) = R \cup C^c$.

Proposition 54. Let $D = H$ be the overlapping index. For this D, and each $E \in \mathscr{B}$, $\exists F(E) \notin \mathscr{A}$ such that

$$\forall x \notin F(E): \rho_x = \rho_x(D, E) = +\infty$$

Proof. Let $\rho = \sup_{x \in C(E)} \rho_x$. It follows from Proposition 53 that $\exists F \notin \mathscr{A}$:

$$\forall y \notin F: \rho_y = \rho \tag{5}$$

If $\rho = +\infty$, there is nothing more to prove. Suppose $\rho < +\infty$. Let O be the overlapping core and $x \in O \cap C \cap F^c$. Then $\rho_x = \rho$. By the proof of Proposition 53, $\forall b \geqslant 1$, $\exists \{n_{l,b}\}$ with $\rho(\{n_l, b\}) = \rho$ and such that

$$\begin{aligned} \forall y \in C - R_b&: \lim_l P^{(n_{l,b} - b)}(y, E) = \bar{P}(E) \\ \forall y \in C - R_{b+D}&: \lim_l P^{(n_{l,b+D} - b - D)}(y, E) = \bar{P}(E) \end{aligned} \tag{6}$$

where

$$P^{(b)}(x, C - R_b) = 1, \qquad P^{(b+D)}(x, C - R_{b+D}) = 1$$

Since $x \in O$, this implies by Proposition 43.1 that $\exists b$:

$$A = (C - R_b) \cap (C - R_{b+D}) \in \mathscr{A}$$

Let $y \in A$; the fact that $\rho(\{n_{l,b+D-b-D}\}) = \rho$ together with the second relation in Eq. (6) shows that $\rho_y \geqslant \rho(\{n_{l,b+D-b-D}\}) \geqslant \rho + 1$. This contradicts Eq. (5) and so $\rho = +\infty$.

Proposition 55. For each $E \in \mathscr{B}$, $\exists A = A(E) \in \mathscr{A}$ and $\{m_i\}\uparrow$ such that $\rho(\{m_i\}) = +\infty$ and $\forall \varepsilon > 0$, $\exists i_0(\varepsilon)$:

$$\forall y \in A, i \geqslant i_0(\varepsilon): P^{(m_i)}(y, E) \geqslant \bar{P}(E) - \varepsilon \tag{7}$$

Proof. By Proposition 54, we may choose the x in the proof of Proposition 53, such that $\rho_x = +\infty$. Since $C - R \in \mathscr{A}$ and

$$C - R = \bigcup_b \bigcup_l (C - R_{l,b})$$

$\exists l$ and b:

$$C - R_{l,b} \in \mathscr{A}$$

Take $A = C - R_{l,b}$, $m_l = n_{l,b-b}$, $i_0 \geqslant l$ and $2^{-i_0} < \varepsilon$, then Eq. (7) is true by Eq. (3).

Proposition 56. Suppose $D = 1$ and that

$$(H)\colon\ E_n \in B,\ E_n \downarrow 0 \Rightarrow \bar{P}(E_n) \downarrow 0$$

Then for each $E \in \mathscr{B}$, $\exists F(E) \notin \mathscr{A}$.

$$\forall x \notin F(E)\colon\ \lim_n P^{(n)}(x, E) = \bar{P}(E) \tag{8}$$

Proof. Let E and $\varepsilon > 0$ be given, and A be the set in Proposition 55. Then $A^\infty \neq \varnothing$ by Proposition 19. For $n \geqslant 1$ we define

$$A_n = \left\{ y \in A^\infty \,\middle|\, \sum_{j=1}^{n} K^{(j)}(y, A) > 1 - \varepsilon \right\}$$

Thus, $A_n \uparrow A^\infty$. Hence by (H), $\exists k$;

$$\bar{P}(A^\infty - A_k) < \varepsilon$$

By Proposition 52, $\exists F_1 = F_1(\varepsilon) \notin \mathscr{A}$:

$$\forall x \notin F_1 : \overline{\lim_n}\, P^{(n)}(x, A^\infty - A_k) = \bar{P}(A^\infty - A_k) < \varepsilon$$

Hence if $x \notin F_1$, $\exists m_0(x)$;

$$\forall m \geqslant m_0(x)\colon P^{(m)}(x, A^\infty - A_k) < \infty$$

Thus we have

$$\forall x \notin F_1,\ m \geqslant m_0(x)\colon\ P^{(m)}(x, A_k) > 1 - \varepsilon \tag{9}$$

Let $\{m_i\}$ be the sequence in Proposition 55. Since $\rho(\{m_i\}) = +\infty$, we may suppose, by taking a subsequence, that it is of the form $\{m_\nu + j;\ 0 \leqslant j \leqslant k - 1,\ \nu \geqslant 1\}$. We have by Proposition 55,

$\forall z \in A,\ \nu \geqslant \nu_0(\varepsilon)$:

$$P^{(m_\nu + j)}(z, E) \geqslant \bar{P}(E) - \varepsilon, \qquad 0 \leqslant j \leqslant k - 1 \tag{10}$$

Let $x \in A^\infty - F_1$ and

$$n > m_0(x) + k + m_{\nu_0(\varepsilon)}$$

Then $\exists \nu = \nu(n) \geqslant \nu_0(\varepsilon)$ such that $m_\nu < n - m_0(x) - k \leqslant m_{\nu+1}$. Put $n - m_\nu - k = m > m_0(x)$. We have

$$P^{(n)}(x, E) \geqslant \int_{A_k} P^{(m_\nu + k)}(y, E)\, P^{(m)}(x, dy) \tag{11}$$

For $y \in A_k$, and $\nu \geqslant \nu_0(\varepsilon)$:

$$P^{(m_\nu+k)}(y, E) \geqslant \sum_{j=1}^{k} K^{(j)}(y, A) \inf_{z \in A} P^{(m_\nu+k-j)}(z, E)$$
$$\geqslant (1-\varepsilon)(\bar{P}(E)-\varepsilon) \tag{12}$$

by the definition of A_k and Eq. (10). Hence by Eqs. (11), (12), and (9):

$$P^{(n)}(x, E) \geqslant (1-\varepsilon)(\bar{P}(E)-\varepsilon)\, P^{(m)}(x, A_k) \geqslant (1-\varepsilon)^2(\bar{P}(E)-\varepsilon)$$

Thus

$$\forall x \in A^\infty - F_1(\varepsilon): \varliminf_n P^{(n)}(x, E) \geqslant (1-\varepsilon)^2(\bar{P}(E)-\varepsilon)$$

and consequently, if $F_2 = \bigcup_{l=1}^{\infty} F_1(1/l)$,

$$\forall x \in A^\infty - F_2: \varliminf_n P^{(n)}(x, E) \geqslant \bar{P}(\varepsilon)$$

By Proposition 52, $\exists F_3 \notin \mathscr{A}$ such that

$$\forall x \in X - F_3: \varlimsup_n P^{(n)}(x, E) \leqslant \bar{P}(E)$$

Therefore

$$\forall x \in A^\infty - (F_2 \cup F_3): \lim_n P^{(n)}(x, E) = \bar{P}(E)$$

Since $X - A^\infty \notin \mathscr{A}$, $F_2 \notin \mathscr{A}$, we may take the $F(E)$ asserted in Proposition 56 to be $(X - A^\infty) \cup F_2 \cup F_3$.

Proposition 56.1. Under the hypothesis in Propsotion 56, $\bar{P}(\cdot)$ is a probability measure on $\mathscr{B}$.

Proof. $\forall E \in \mathscr{B}$, $\bar{P}(E) \geqslant 0$ by definition. Next, let $\{E_j, 1 \leqslant j \leqslant l\}$ by disjoint sets in $\mathscr{B}$, E their union, and let $F = F(E) \cup \bigcup_{j=1}^{l} F(E_j)$ where for each E, $F(E)$ is the set given in Eq. (8). If $x \notin F$, we have

$$\bar{P}(E) = \lim_n P^{(n)}(x, E) = \lim_n \sum_{j=1}^{l} P^{(n)}(x, E_j) = \sum_{j=1}^{l} \bar{P}(E_j)$$

Hence $\bar{P}(\cdot)$ on $\mathscr{B}$ is finitely additive, and so it is countably additive by (H). Finally, if $x \notin F(X)$:

$$\bar{P}(X) = \lim_n P^{(n)}(x, X) = 1$$

Therefore, $\bar{P}(\cdot)$ is a probability measure.

Proposition 57. Let $\{I_i,\ 0 \leqslant i \leqslant D-1\}$ be the maximum cycle in X. For each $E \in \mathscr{B}$, $\exists F(E) \notin \mathscr{A}$ such that

$$\forall x \in I_i \backslash F(E): \quad \lim_{\substack{n \to \infty \\ n \equiv j-i (\mathrm{mod}\, D)}} P^{(n)}(x, E) = \bar{P}(E \cap I_j)$$

Proof. Put $Q(x, E) = P^{(D)}(x, E)$, $x \in X$, $E \in \mathscr{B}$ so that $Q^{(n)} = P^{(nD)}$. If $x \in I_i$, then

$$Q^{(n)}(x, I_j) = 1 \qquad \text{for} \quad n \equiv j - i (\mathrm{mod}\, D)$$

by the properties of a cycle. Let Q_j be the restriction of Q to $I_j \times (\mathscr{B} \cap I_j)$. By Proposition 48.1, for any $x \in I_j \cap G \cap G'$: $h(x) = D$. Hence for any consequent sequence $\{C_n,\ n \geqslant 1\}$ of x with respect to Q_j, we have $C_n \cap C_{n+1} \in \mathscr{A}$ for some n. This means the overlapping index of I_j with respect to Q_j is equal to 1 and Proposition 56 is applicable. If $x \in I_j$ and $E \subset I_j$, then $P^{(n)}(x, E) = 0$ unless $D \mid n$; it follows that for a suitable x:

$$\bar{Q}_j(E) = \overline{\lim_n}\, Q_j^{(n)}(x, E) = \overline{\lim_n}\, P_j^{(n)}(x, E) = \bar{P}(E)$$

Namely, the $\bar{Q}_j$ corresponding to Q_j coincides with $\bar{P}$ restricted to $\mathscr{B} \cap I_j$. Therefore by Proposition 56, $\exists F_j \subset I_j,\ F_j \notin \mathscr{A}$ such that if $x \in I_j - F_j$ and $E \in \mathscr{B} \cap I_j$, then

$$\lim_n Q_j^{(n)}(x, E) = \bar{Q}_j(E) = \bar{P}(E)$$

For an arbitrary E, we write $E = \bigcup_{j=0}^{D-1} (E \cap I_j)$; if $x \in I_i$, then

$$P^{(n)}(x, E) = P^{(n)}(x, E \cap I_j) \qquad \text{for} \quad n \equiv j - i (\mathrm{mod}\, D)$$

The general result in Proposition 57 follows.

Proposition 57.1. Let $\bar{P}_j$ be the restriction of $\bar{P}$ to I_j, $0 \leqslant j \leqslant D-1$. Then for each j, $\bar{P}_j(\cdot)/\bar{P}_j(I_j)$ is a probability measure on $\mathscr{B} \cap I_j$.

Proposition 58. [Under (H)] For each cl. set C we have $\bar{P}(C) = 1$. For each abs. ess. set A we have $\bar{P}(A) > 0$.

Proof. If C is cl. then $F(C)^c \cap C \neq \varnothing$ where $F(C)$ is the set in Proposition 52 corresponding to C. Take any x in $F(C)^c \cap C$, we have

$$\bar{P}(C) = \overline{\lim_n}\, P^{(n)}(x, C) = \overline{\lim_n}\, 1 = 1$$

Next for $A \in \mathscr{A}$, let

$$A_k = \left\{ x \in A^\infty \,\middle|\, \sum_{j=1}^{k} K^{(j)}(x, A) > \frac{1}{2} \right\}$$

Since $A_k \uparrow A^\infty$, it follows from (H) that $\bar{P}(A^\infty - A_k) \downarrow 0$ as $k \uparrow \infty$. Since $\bar{P}(\cdot)$ is finite subadditive by definition and $\bar{P}(A^\infty) = 1$ by what has just been proved, we have $\bar{P}(A_k) > \frac{1}{2}$ for all sufficiently large k. Fix such a k here. We have for every $n \geqslant 1$, $j \geqslant 1$:

$$P^{(n+j)}(x, A) \geqslant \int_{A_k} K^{(j)}(y, A)\, P^{(n)}(x, dy)$$

$$\sum_{j=1}^{k} P^{(n+j)}(x, A) \geqslant \int_{A_k} \sum_{j=1}^{k} K^{(j)}(y, A)\, P^{(n)}(x, dy) > \tfrac{1}{2} P^{(n)}(x, A_k)$$

Let $x \in F(A)^c \cap F(A_k)^c$. Then we have by Proposition 52:

$$\overline{\lim_n}\, P^{(n)}(x, A) = \bar{P}(A), \qquad \overline{\lim_n}\, P^{(n)}(x, A_k) = \bar{P}(A_k)$$

Hence

$$\begin{aligned} k\bar{P}(A) &= \sum_{j=1}^{k} \overline{\lim_n}\, P^{(n+j)}(x, A) \geqslant \overline{\lim_n} \sum_{j=1}^{k} P^{(n+j)}(x, A) \\ &\geqslant \tfrac{1}{2} \overline{\lim_n}\, P^{(n)}(x, A_k) = \tfrac{1}{2} \bar{P}(A_k) > \tfrac{1}{4} \end{aligned}$$

Thus

$$\bar{P}(A) > \frac{1}{4k} > 0$$

REFERENCES

1. Chung, K. L. (1964). The general theory of Markov processes according to Doeblin. Z. Wahrscheinlichkeitstheorie 2, 230–254. See also Errata, loc. cit. 12, 172 (1969). On p. 239, line 9, insert "or empty" at end; on p. 241, line 7, read "indecomp." for "decomp."
2. Doeblin, W. (1940). Éléments d'une théorie générale des chaînes simples constantes de Markoff. *Ann. Sci. École Norm. Sup. (3)*, 57 (2), 61–11.

SELECTED WORKS OF
KAI LAI CHUNG

This unique volume presents a collection of the extensive journal publications written by Kai Lai Chung over a span of 70-odd years. It is produced to celebrate his 90th birthday. The selection is only a subset of the many contributions that he has made throughout his prolific career. Another volume, *Chance and Choice*, published by World Scientific in 2004, contains yet another subset, with four articles in common with this volume. Kai Lai Chung's research contributions have had a major influence on several areas in probability. Among his most significant works are those related to sums of independent random variables, Markov chains, time reversal of Markov processes, probabilistic potential theory, Brownian excursions, and gauge theorems for the Schrödinger equation.

As Kai Lai Chung's contributions spawned critical new developments, this volume also contains retrospective and perspective views provided by collaborators and other authors who themselves advanced the areas of probability and mathematics.

World Scientific
www.worldscientific.com
6976 hc

ISBN-13 978-981-283-385-3
ISBN-10 981-283-385-4

9 789812 833853